Life in Subsurface Pools: Insight into Microbial Diversity and Dynamics in the Endokarst Environment

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der Universität Zürich

von

Tatiana Shabarova

aus Russland

Promotionskomitee

Prof. Dr. Jakob Pernthaler (Vorsitz)

Prof. Dr. Leo Eberl

Dr. Franco Widmer

Zürich, 2013

«Думать это не удовольствие, это обязанность.»

Аркадий и Борис Стругацкие «Улитка на склоне»

Dedicated to Silvia Schmassmann

Table of contents

Abbreviations

DOC	dissolved organic carbon
FT-IRC-MS	Fourier-transform ion cyclotron resonance-mass spectrometry
HRH	Caving Association of the Hohgant Region
ISSKA	Institute of Swiss Speleology and Karstology
MCD	mean community distance
NGS	next generation sequencing
NRI	net relatedness index
NTI	nearest taxon index
OTU	operational taxonomic unit
PCA	polycyclic aromatic
rRNA	ribosomal ribonucleic acid
SGH	Swiss Speleological Society
TOC	total organic carbon
T-RFLP	terminal restriction fragment length polymorphism
WHO	World Health Organization

Summary

The term “karst” describes a special type of landscape featuring caves and extensive underground water systems that develop primarily on soluble rocks. Subterranean karst is an important environment for the conduction and storage of water of atmospheric origin. Roughly 20-25% of the global population largely or entirely depends on karst water resources, which are coming under increasing anthropogenic pressure and are in great need of protection and sustainable management. To date, little is known about the microbial communities of karst water and their transformation during their passage through karst systems. Caves are one important type of subsurface karst formation. They are per definition accessible for human exploration and contain all karst components. Therefore, they present a unique opportunity for studies, modeling and prediction of the behavior of various kinds of karst water, and for the investigation of the microbial communities within. Subsurface water pools are ubiquitous in vadose (unsaturated) and epiphreatic (periodically flooded) zones and represent a special type of aquatic karst environment with different stability. Bärenschacht cave located in Bernese Oberland is a profound karst system consisting of a vertical entrance part and a large labyrinth. Its lower realms harbour an extended epiphreatic zone with numerous small ponds. At the same time, plenty of vadose pools located in the upper part of the labyrinth do not have contact with ground water but are instead fed by dripping water. Two vadose and three epiphreatic ponds from this cave system were chosen as research sites to assess the diversity of microbial communities in endokarst water and to observe the transformations of bacterial assemblages after flood events.

The studied pools from zones with different water saturation had contrasting chemical and microbiological parameters. The vadose pools displayed relatively high conductivity, distinct ion composition and low bacterial abundances ($< 10^4$ cells ml^{-1}). By contrast, conductivity in the epiphreatic pools did not exceed 250 µS cm^{-1}, and bacterial numbers ranged between 1×10^5 and 5×10^5 cells ml^{-1}. Unexpectedly high microbial phylotype richness was found, even in the two vadose pools with scant water exchange and low cell numbers. Most of detected operational taxonomic units (OTUs), especially in the vadose ponds, were not associated with typical freshwater bacterial linages. Although β-*Proteobacteria* were the dominant group in all three investigated pools, the vadose systems showed only minimal overlap of phylotypes not only with epiphreatic ponds but also with each other. Some distinct phylogenetic groups were independently present in different systems, e.g. members of candidate phylum OP3 were observed exclusively in the vadose pool with high nitrate concentration. We conclude that particular granular aquifers feeding such pools might determine both, chemistry and microbial community composition in these habitats.

In a follow-up study, high similarity of microbial community composition and reproducible transformations of bacterial assemblages during water residence after flood events were detected in the epiphreatic pools, as assessed by terminal restriction fragment length polymorphism (T-RFLP) of 16S rRNA genes, microscopy and sequence analysis (by clone libraries and next generation sequencing, NGS). A seasonal recurrence of the microbial assemblages in the epiphreatic pools was demonstrated by comparison of community composition between samplings taken in two consecutive winters. The families *Methylophilaceae* and *Comamonadaceae* of β-*Proteobacteria* represented the two most abundant linages, with high potential for persistence. A large proportion of OTUs primarily affiliated with other bacterial linages was present only in samples collected shortly after flood events, whereas only few genotypes were favored by the pool environment. A small proportion of all OTUs (74 of 4804) was detected in all samples analyzed by NGS. This core of microbial community harboured more than half of all sequences obtained by NGS. The nearest neighbors of more than two-thirds of the core OTUs in a curated database of 16S rRNA genes (SILVA) were microorganisms from various aquatic habitats (including karst water). Reproducible temporal patterns of community development were reflected in the relative abundances changes of OTUs related to β-*Proteobacteria* and several other phylogenetic groups. By contrast, some OTUs affiliated with *Bacteroidetes* displayed a much stronger connection to the different floods then to residence time. Water samples originating from different floodings could be distinguished according to the composition of dissolved organic carbon (DOC), as analyzed by Fourier-transform ion cyclotron resonance-mass spectrometry (FT-IRC-MS). In addition, FT-IRC-MS also revealed a transformation of DOC characteristics over time. Surprisingly, the influence of deterministic community assemblage processes, as assessed by phylogenetic clustering, was detected in all samples, even of the influx water. At the same time, the sample obtained during a flood event did not display the significant clustering of deeply branching phylogeny with any other samples. This confirmed our hypothesis that the bacterial inocula of the ponds originated from various distinct surface and subsurface niches, each one with it own selective properties. The reproducible decrease in bacterial diversity and increasing relatedness at a lower phylogenetical level with residence time reflected the importance of environmental filtering in the epiphreatic habitats.

Zusammenfassung

Der Begriff „Karst“ bezeichnet eine besondere Form von Landschaft, welche durch ein gut entwickeltes unterirdisches Drainagesystem und die Bildung von Höhlen charakterisiert ist. Karst entwickelt sich vor allem in gut löslichen Gesteinstypen. Subterraner Karst ist wichtig für das Leiten und Speichern des Niederschlagswassers. Ca. 20 bis 25% der Weltbevölkerung sind zum großen Teil oder vollständig auf Karstwasserressourcen angewiesen. Diese geraten zunehmend unter den Druck anthropogener Nutzung und benötigen daher dringend Schutz und nachhaltiges Management. Zur Zeit ist wenig über die mikrobiellen Gemeinschaften des Karstwassers und deren Veränderungen während ihrer Passage durch das karstifizierte Gestein bekannt. Höhlen sind eine wichtige Formation des unterirdischen Karstes. Sie sind per Definition für Forscher zugänglich und enthalten alle Karstformen. Daher bieten sie eine einzigartige Gelegenheit für die Untersuchung, die Modellierung und die Vorhersage des Verhaltens verschiedener Arten von Karstwasser sowie für die Erforschung seiner mikrobiellen Gemeinschaften. In den vadosen (ungesättigte) und epiphreatischen (periodisch gefluteten) Zonen des Karstes finden sich häufig Wasserbecken und –tümpel (Pools), die eine besondere Art aquatischen Lebensraumes mit unterschiedlicher Stabilität repräsentieren. Der Bärenschacht ist eine tiefe Karsthöhle im Berner Oberland. Sie beginnt mit einem vertikalen Eingangsbereich, welcher in ein grosses Labyrinth führt. In seinem unteren Teil verbirgt sich eine ausgedehnte epiphreatische Zone mit zahlreichen Pools. Gleichzeitig finden sich viele vadosen Pools im oberen Teil des Labyrinthes, die ohne jeglichen Kontakt zum Grundwasser nur vom Tropfwasser gespeist werden. In dieser Arbeit wurden zwei vadose und drei epiphreatische Pools aus dem Bärenschacht ausgewählt, um die Diversität der mikrobiellen Gemeinschaften im Endokarst-wasser und deren Transformation nach Hochwasserereignissen zu untersuchen.

Die untersuchten Pools zeigten abhängig von der Wassersättigung der entsprechenden Zone charakteristische chemische und mikrobiologische Parameter. Die vadosen Teiche zeichneten sich durch relativ hohen Leitfähigkeit, unterschiedliche Ionenzusammensetzung und tiefen bakteriellen Dichten ($<10^4$ Zellen ml^{-1}) aus. Dagegen überschritt die Leitfähigkeit in den epiphreatischen Pools nie 250 µS cm^{-1}, und die bakterielle Abundanz in diesen Habitaten variierte zwischen 1×10^5 und 5×10^5 Zellen ml^{-1}. Es wurde eine unerwartet hohe Anzahl verschiedener mikrobiellen „Arten“ in den Pools festgestellt, auch in den beiden vadosen mit vernachlässigbarem Wasseraustausch und geringen Zellzahlen. Die meisten der detektierten Genotypen (OTUs, „Operational Taxonomic Units“), insbesondere aus den vadosen Pools waren nicht nahe mit typischen Süßwasserbakterien verwandt.

Obwohl in allen untersuchten Tümpeln β-*Proteobacteria* als dominierende Bakteriengruppe gefunden wurde, zeigten die vadosen Systeme nur eine geringe Anzahl gemeinsamer Phylotypen sowohl mit den epiphreatischen Habitaten als auch untereinander. Einige der oft exotischen phylogenetischen Gruppen wurden ausschließlich in einem der Systeme nachgewiesen, z.B. Verwandte des „candidate phylums" OP3 nur in dem vadosen Habitat mit hoher Nitratkonzentration. Daraus lässt sich schliessen, dass die einzelnen granulären Wasseradern, welche solche Pools speisen, sowohl deren Chemie, als auch die jeweilige Zusammensetzung der mikrobiellen Gemeinschaften in diesen Lebensräumen bestimmen könnten.

Mit der Hilfe von unterschiedlichen Methoden (unter anderem Restriktionsfragment-Längenpolymorphismus von 16S rRNS Genen, Mikroskopie und vergleichende Sequenz-Analyse) wurde eine große Ähnlichkeit der Zusammensetzung der bakteriellen Gemeinschaften in den epiphreatischen Pools festgestellt und deren reproduzierbare zeitliche Transformation nach Hochwasserereignissen nachgewiesen. Das saisonale Wiederauftreten vieler Phylotypen wurde durch Vergleich der Artenzusammensetzung von Proben aus zwei aufeinander folgenden Wintern demonstriert. Die zwei am häufigsten beobachteten Familien waren *Methylophilaceae* und *Comamonadaceae* der β-*Proteobacteria;* diese wiesen auch hohe Persistenz in den Systemen auf. Der Großteil der Phylotypen aus anderen bakteriellen Gruppen konnten nur in Proben während oder kurz nach dem Hochwasser nachgewiesen werden, und nur wenige Genotypen wurden durch die Umgebung begünstigt. Ein kleiner Anteil von OTUs (74 von insgesamt 4804) war in allen analysierten Proben präsent. Dieser mikrobiellen Kerngemeinschaft umfasste allerdings mehr als die Hälfte aller produzierten Sequenzen, und die nächsten Verwandten von über zwei Drittel dieser Kern-OTUs waren Mikroorganismen aus verschiedenen aquatischen Habitaten (einschließend Karstwasser). Die Analyse der relativen Häufigkeiten von β-*Proteobacteria* und mehreren anderen phylogenetischen Gruppen ermöglichte die Gruppierung von Proben in Abhängigkeit von der Aufenthaltszeit des Wasser nach Flutereignissen. Im Gegensatz dazu zeigten einige Phylotypen der *Bacteroidetes* eine viel deutlichere Verbindung zu den unterschiedlichen Flutereignissen. Wasserproben aus verschiedlichen Überflutungsereignissen konnten auch anhand der Zusammensetzung des gelösten organischen Kohlenstoffes (detektiert mit Hilfe der Fourier-transformierten Ionenzyklotron-Resonanz-Massenspektrometrie) aufgetrennt werden. Darüber hinaus wurden auch Veränderungen der Eigenschaften des gelösten organischen Kohlenstoffs mit der Wasseraufenthaltszeit in den Pools beobachtet. Überraschenderweise wurde ein Einfluss von deterministischen Prozessen auf die Zusammensetzung der bakteriellen Gesellschaften anhand ihrer signifikant geklumpten phylogenetischen Struktur auch in einer Probe welche während des Hochwassers gesammelt wurde, nachgewiesen. Hingegen

zeigte diese Probe in ihrer tiefen phylogenetischen Struktur keine signifikante „Clusterisierung“ mit den anderen Proben. Diese Beobachtung spricht für die Hypothese, dass das bakteriellen Inokulum des Flutwassers aus mehreren verschiedenen ober- und unterirdischen Nischen stammt, welche jede für eine bestimmte Gemeinschaft selektiv ist. Die reproduzierbare Abnahme der bakteriellen Vielfalt und das steigende „Clusterisierung“ der tiefen Phylogenien mit zunehmender Wasseraufenthaltszeit zeigt ausserdem die Bedeutung der Habitatsselektivität in den untersuchten epiphreatischen Pools.

1. Introduction

1.1 Carbonate rocks and karst: Definition and importance

There are plenty of reasons for the attractiveness of carbonate rocks for scientists. On the one hand carbonates are of highly diversified origin and represent the most important source of geological, fossil and climatic records. At the same time, this rock type contains at least 40% of the global hydrocarbon resources, plays a role in the carbon cycle and hosts metal and other deposits. Carbonates also represent ever-developing, changing and renewable systems due to the incessant sedimentation from the Precambrian age until modern times, and the ongoing dissolution and remineralization processes (Tucker 1990).

Carbonate rocks typically undergo karstification, a solution derived process acting also on rocks such as gypsum or halite. In fact, karst can be defined as a landscape with distinctive hydrology and landforms that developed due to high rock solubility and secondary porosity. The term “karst” is a Germanisation of the word Carsus or Carso used to identify the karstic landscape shared between Slovenia and Italy. The origin of these names can be traced to Indo-European roots, the words “kara” vs. “gara” meaning stone (Kranjc 2001). Karst can be divided in exo- (surface) and endo- (subsurface) karst. The most widespread formation of exokarst are dolines (sinkholes), karren and hums (towers) (Fig. 1).

Figure 1: *(a) Innerbergli karrenfield in Bernese Oberland, photo by D. Sanz; (b) Kavakuna doline in Papua New Guinea, photo by S. Avarez*

Large ice-free areas of the Earth that form more than 10% of the land surface are underlaid by karst, 90% of which is represented by karst formations on carbonate rocks (Ford and Williams 2007). The well-developed drainage system that is characteristic for this landscape is responsible for its high water conduction and storage capacity. Groundwater originating from karst is, therefore, one of the most important sources of drinking water. Around 20-25% of the global population depend on water obtained from karst drainage and storage systems. In Europe, carbonate terrains occupy 35% of the land surface, and karst water is the only available freshwater source in some regions. In the context of the decreasing glacier water stores, karst water resources are coming under increasing pressure

and are in great need of conservation, sustainable management and research (COST-65 1995, Ford and Williams 2007).

1.2 The karstification process

Conventional karstification results mostly from a mineral dissolution by water. However rock solubility is an essential but not sufficient geological precondition for karst development, as low primary porosity of rock is also required. The intensity and velocity of this process are dependent on climatic parameters; karst formation is most rapid in humid and warm conditions and correlates linearly with the amount of runoff (Priesnitz 1974). The solubility is mineral-specific. The most widespread carbonate rock types undergoing karstification are limestone, $CaCO_3$ (minerals calcite and aragonite), and dolomite, $CaMg(CO_3)_2$. The differences in their solubility are given in Table 1, which also shows the dependency of solubility on the partial pressure of CO_2 (p_{CO_2}).

Table 1: Solubility of calcite and dolomite at different P_{CO_2} (Freeze and Cherry 1979)

Mineral	*Solubility 1 (mg l^{-1})*	*Solubility 2 (mg l^{-1})*
Calcite	60	400
Dolomite	50	300

Solubility 1 at $p_{CO_2}=10^{-3}$ bar, Solubility 2 at $P_{CO_2}=10^{-1}$

The increase of carbonate solubility at higher CO_2 can be readily explained with equations (1)-(3):

$$CO_2\,(g) \leftrightarrow CO_2\,(aq) \quad (1)$$

$$CO_2\,(aq) + H_2O \leftrightarrow H_2CO_3^{\,0} \quad (2)$$

$$CaCO_3\,(s) + H_2CO_3^{\,0}(aq) \leftrightarrow Ca^{2+} + 2HCO_3^{\,-} \quad (3)$$

The third equation is the partial reaction, which together with the two additional equations (4) and (5) describes calcite dissolution (Plummer et al 1978). This process can also be summarized in one equation (6).

$$CaCO_3\,(s) + H^{+} \leftrightarrow Ca^{2+} + HCO_3^{\,-} \quad (4)$$

$$CaCO_3\,(s) + H_2O \leftrightarrow Ca^{2+} + HCO_3^{\,-} + OH^{-} \quad (5)$$

$$CaCO_3\,(s) + CO_2\,(aq) + H_2O \leftrightarrow Ca^{2+} + 2HCO_3^{\,-} \quad (6)$$

According to the depicted reactions (4)-(6), the concentration of CO_2 in water and pH are two important factors influencing the dissolution kinetic. In systems that are at equilibrium

with the atmosphere, the concentration of CO_2 depends on p_{CO_2} (whose actual value at sea level is ca. 39 Pa, corresponding to the concentration of 0.039%) and temperature (the higher the temperature the lower the solubility of CO_2). Thus, for karstification developing on bare rocks lacking soil and vegetation, a temperate climate would provide the best conditions. The situation is different for covered karst, where biological processes in the soil layer rather than the atmosphere are the most relevant source of CO_2. At such conditions, temperature together with humidity and primary production levels are the factors determining soil CO_2 concentration and fluxes (Lloyd and Taylor 1994, Raich and Schlesinger 1992, Rustad and Fernandez 1998). In tropical regions, the proportion of CO_2 in soil gases normally ranges between 0.2 and 11%, but there are some reports of up to 17% (Derbyshire 1976, Hashimoto et al 2004, Köhler 2009, Kursar 1989, Liu et al 2010). In soils of moderate climate, the usual range is between 0.1 and 3.5% but values up to 10.2% are occasionally reported (Bekele et al 2007, Davidson et al 2007, Derbyshire 1976, Jassal et al 2005). Thus, plant and especially microbiological respiration is able to increase the CO_2 concentration by more than hundredfold compared to the atmospheric conditions (Gabrovsek et al 2000) and to accelerate karstification. Acidic microbial exudates also increase solubility of minerals. Moreover, the dissolution kinetic is also affected by other less significant factors such as rock granulation, which will not be discussed here.

The above described conventional karstification based mostly on water and carbonic acid dissolution result in the most common type of karst. Along with it other karst forms exist. In barokarst karstification is induced by high pressure corrosion, and thermal springs and steam play a key role in thermokarst formation. Another special karst type is hypokarst, where karstification starts from the inside of the bedrock; this often happens at the border of oxic and anoxic zones in the presence of H_2S rich ground water. The karstification mechanism in such cases is driven by sulfuric acid dissolution. The mechanism of sulfuric acid karstification was firstly proposed in the 1970's by Egemeier (Egemeier 1981). In this model H_2S volatizes into the cave atmosphere and undergoes oxidation to sulfuric acid on wet bedrock surfaces, causing replacement of the carbonate by gypsum. Gypsum can be easily dissolved in water that considerable increases void volume. Other dissolution processes occurring at or just below the water surface have also been attributed to sulphuric acid karstification (Hill 1990, Palmer 1991). These mechanisms explain the genesis of some famous karst caves, e.g., Lechuguilla (USA), Frasassi (Italy), Novaya Afonskaya (Georgia). At study sites with the above mentioned karst types, enrichment of light isotopes of sulphur have been detected in gypsum deposits (Hill 1990, Klimchouk et al 2000), hinting at a possible involvement of microbes in such processes. While some aspects of the microbial role in this kind of karstification are still under discussion, Engel and coworkers have shown that the activity of sulfur-oxidizing microorganisms affiliated with β-, γ-, and ε-*Proteobateria*

may contribute more to the H_2S oxidation and dissolution of carbonates than abiotic processes under certain conditions (Engel et al 2004). Specifically, the utilization of H_2S already in the subaquatic environment and local sulfuric acid production minimizes the amount of volatized H_2S and its ability to dissolve carbonate.

1.3 Karst water

Water is the factor driving karstification and its special behaviour is an important attribute in developed karst. For a better understanding of the processes occurring in karst water, it is important to distinguish between the following karst zones according to water persistence:

I. Unsaturated (or vadose) zone above the water table, where voids in the rock are only partially occupied by water. Water transport is mostly vertical.
II. Saturated (or phreatic) zone below the water table containing ground water. Water transport is mostly horizontal.
III. Epiphreatic zone, defined as a zone which is sometimes (after heavy rain or during snow melting time) located below the water table.

Rock porosity is another important factor influencing many characteristics of the behaviour of karst aquifers. Carbonate rocks may acquire a range of porosity that affects their capacity to store and transmit water. They are, therefore, commonly differentiated into three types, according to the nature of the voids in which the water is stored and through which it is transmitted:

I. Granular (or matrix, water moves through bulk rock pores).
II. Fracture (water moves through voids arising from rock folding and faulting).
III. Conduit (via pipes, develops by circulating groundwater along penetrable fissures).

In practice, most karst aquifers have components of each type, but by definition they must possess a significant conduit component (Ford et al 2007).

The chemical parameters of karst water can vary significantly but can nevertheless be defined within particular ranges. The common ions in karst water are the products of calcite and dolomite dissolution: Ca^{2+}, Mg^{2+} and carbonate ions. The conductivity of water in unsaturated karst usually ranges between 150 and 400 µS cm^{-1}. While the water from porous aquifers with relatively long residence time shows conductivity values around the top of this range, conduit and fissured aquifers are associated with low conductivity due to the high velocity of water flow and short residence times (T.S., personal observation, Bakalowicz and Mangin 1980). Conductivity values higher than this range (at normal CO_2 levels) often

indicate the presence of additional ions, particularly of sulphate (which typically originates from gypsum dissolution), or of sodium, potassium, and chloride, the latter mostly resulting from halite (NaCl) and sulvite (KCl) dissolution (Ford and Williams 2007). The solubility of these minerals is much higher than of carbonates and can reach 2400 mg l^{-1} for gypsum and 360 g l^{-1} for halite. The presence of different metal deposits in carbonates can also influence the ionic composition of the karst water.

While most mineral ions are the result of bedrock dissolution, compounds containing nitrogen, phosphorus and organic carbon species are typically of surface, e. g. soil, origin, and can be used as natural tracers of infiltration (Baker et al 1997, Batiot et al 2003). The concentrations of these substances evolve in a way inverse to that of the rock-originating ionic components, with increases during high water flow situations (short residence time) and decreases in periods of low water flow (long residence time) (Xuan et al 2011). The following nitrogen sources can be distinguished: rain water, surface runoff, rock matrix (which can store and release compounds), rare nitrogenous minerals (Boyce et al 1976) and other pollutants. Artificial contamination, such as agriculture, use of soil fertilisers and industry have increased the concentration of nitrogen species in ground water during the second half of the 20th century in Nord America, Europa and Asia (Drake and Bauder 2005, Fuhrer et al 1999, Hansen et al 2011, Nishio 2002). The natural concentrations of nitrate, the most abundant nitrogen compound in karst water, is very difficult to estimate because of scant monitoring data for uncontaminated karst regions; in some publications a range between 0.1 and 10 mg l^{-1} has been defined (Feth 1966), or it is assumed to be <2 mg l^{-1} (Mueller and Helsel 1996). Contaminated karst water aquifers often show nitrate concentrations >50 mg l^{-1} (i.e., the drinking water threshold value (WHO 2011)), and it can accumulate up to 300 mg l^{-1} in ground water (Zhang et al 1996). There are numerous case studies where a proper landscape management has led to a successful decrease in the concentrations of nitrogen species in karst ground water (van Beynen 2011) but it still remains a problem in many regions. The concentrations of phosphorus compounds in karst water have not been in the focus of research until recently. The ability of carbonate rich soils and carbonate rocks to adsorb and precipitate these compounds (Rhue et al 2006) renders the contamination problem less urgent than in the case of nitrate. At the same time, the high differences between drinking water (1 mg kg^{-1} tolerance value in Switzerland) and ecologically relevant thresholds for phosphorus, together with relatively insensitive standard determination methods, likely do not allow to estimate the real extent of the problem. In covered karst, soil is also the most important source of dissolved organic carbon (DOC) to the karst water systems. DOC originates from decomposition of the organic material in the upper soil horizon (Brooks et al 1999) and is the key parameter determining heterotrophic bacterial activity (Culver 1985). DOC concentrations in soils depend on soil properties, vegetation type,

hydraulic characteristics and the thickness and water saturation of the soil horizon. They usually range between 10 and 100 mg l^{-1}, but selectively decrease with soil depth (Drever 1997, van Beynen et al 2002). In forest ecosystems the DOC flux between 0.2 and 1 m depth can be as high as 690 kg ha^{-1} yr^{-1} (Neff and Asner 2001). In karst water, DOC concentrations vary between 0.2 and 10 mg l^{-1} but are frequently <1-2 mg l^{-1} (Clark and Fritz 1997, Kendall and McDonnell 1998, Simmleit and Herrmann 1987, Trček 2008, van Beynen et al 2002). In general, DOC does not significantly differ from total organic carbon (TOC) in carbonate aquifers (Bakalowicz 2003, Batiot 2002).

1.4 Caves

Caves are the most typical forms of endokarst. They might be characterized as the most complex of all landforms because of their non-trivial three dimensional patterns inside the bedrock, developed through the action of numerous and changing hydrological, tectonic, biotic and climatic parameters. A frequently used definition for a cave is ‘a natural underground void, large enough to be entered by people’. Caves can cross all zones of water saturation or be hydrological completely or partial inactive (fossil). The hydrology of active caves corresponds to the properties of well-developed karst; it has a dominating conduit component, but granular and fracture components can also be present. Speleogenesis (cave formation) can be caused by all types of karstification. Since caves feature all karst characteristics, they present a unique opportunity for karst studies, e.g., for the modeling and prediction of the behavior of various kinds of karst water, and for the investigation of the microbial communities and processes within.

1.5 Microorganisms in caves

Caves are often situated in carbonate rocks with geological records tracing back up to several millions years, while at the same time they are usually interconnected with different habitats such as soils and surface waters. Therefore, they may both serve as long-term reservoirs for endemic microbes, and at the same time they may be inhabited by allochthonous microorganisms. Despite the typically stable local conditions of parameters such as temperature, humidity, etc. in a particular cave habitat over years, these factors may differ considerably not only between caves but also within the same cave (Engel 2010). For example, physicochemical parameters may depend on the distance from the entrance (influence of sunlight and climate on the surface), on the absence or presence of water and allochthonous energy sources, chemical compounds suitable for autotrophy, and

concentrations of toxic gases such as CO, H_2S (Engel 2007). Thus, caves provide plenty of exclusive microbial habitats and niches.

Research about cave microorganisms has developed hand in hand with the discovery of new cave entrances and passages by speleologists and also with the development of modern molecular technics. Thus the bulk of microbiological explorations in caves has only occurred during the last three decades.

1.5.1 Microbes in sulfidic caves

Most cave habitats are subjected to strong carbon limitation because the flux of energy from photosynthetically produced biomass is typically interrupted. Sulfidic caves provide favorable conditions for the existence of local ecosystems that are relatively independent from sunlight. Such habitats were found in subsurface karst of hypogenic origin. Their productivity can be attributed to sulfur oxidation, comparable to other well described hotspots in the marine environment and continental waters, such as deep-sea hydrothermal vents, solfataras, and sulfur-rich springs (Boston et al 2006, Engel 2007, Frund and Cohen 1992, Huber et al 1996, Ruby et al 1981). In caves these systems are often characterized by water sources rich in H_2S and poor in oxygen, visible microbial formations (Fig. 2) such as free-floating or attached mats of filamentous bacteria (Chen et al 2009, Engel et al 2003, Engel et al 2004, Macalady et al 2008), and sometimes even by macroscopic ecosystems: In Movile Cave (Romania), a rich fauna including 33 endemic species of invertebrates has been described (Sarbu 2000). There is already solid knowledge on the microbial diversity in sulfidic karst caves, albeit mainly limited to the aerobic part of these habitats. Specifically, ε-*Proteobacteria* along with γ-, β-, and α-proteobacterial lineages have been reported as the most frequent sulfur oxidizers (Engel 2007). Moreover, microbial diversity has often been described in context of biochemical and geological parameters. Sulfur oxidation with oxygen but also with nitrate as electron acceptor has been detected in several caves (Mattison et al 1998). Autotrophic production (as estimated by C^{14} bicarbonate assimilation) in Movile Cave (Romania), Frassasi Cave (Italy), the submarine cave Grotta Azzura (Mexico), and the Lower Kane Cave (USA) ranged between 0.1 and 10 µg C mg dry weight^{-1} h^{-1} (Engel et al 2001, Engel 2007, Porter 1999). The values of heterotrophic microbial processes in these systems were several times lower.

Since the karst environment is a perfect buffering system, karst caves harbor a high variety of sulfidic habitats at circumneutral pH. However, acidic habitats have also been described, e.g. on gypsum layers separating bacteria from carbonate rocks (Hose et al 2000, Macalady et al 2007). Most of the microbes detected in acidic systems are affiliated with *Proteobacteria*, *Acidimicrobium*, *Thermoplasmales* and candidate division TM6. Some

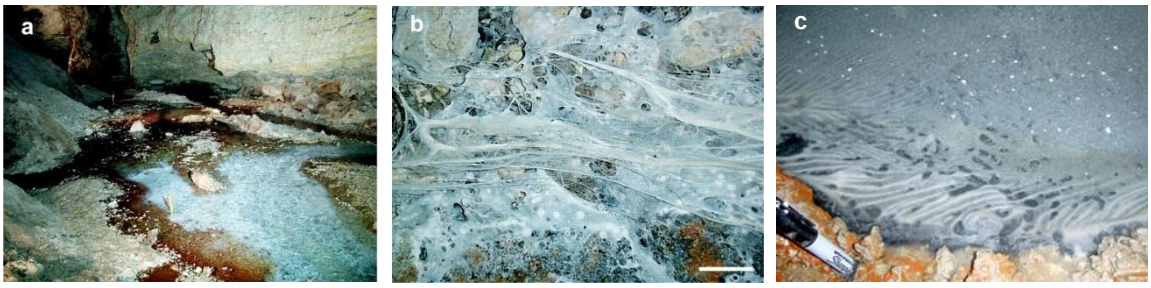


Figure 2: *(a) Photograph of cave passage showing microbial mats growing in sulfidic stream channel formed downstream of the Upper Spring orifice in Lower Kane Cave. The stake in the center of the view is approximately 25 cm high (b) Filamentous microbial mats in cave stream. Scale bar, 10 cm (Engel et al 2003); (c) microbial mat from Frassasi cave, J. Macalady*

studies attempt to connect the microbial activity in such habitats with particular geologic processes (Barton and Luiszer 2005, Engel et al 2004) and to link functional and phylogenetic diversity to nutrient spiraling (Engel et al 2010). The latter study has also provided the first insight into the anaerobic part of the sulfuric system in Lower Kane Cave (USA), dominated by δ-*Proteobacteria* and members of the *Chloroflexii*. Some groups of autotrophic, formatotrophic and acetotrophic sulfate reducers could be identified and linked to shifts of isotopic ratios of C (Engel et al 2010). Among sulfur based microbial metabolic processes, a study on functional genes conducted in Movile Cave (Romania) revealed some contribution of nitrite and ammonia oxidizers to the primary production (Chen et al 2009). Relatively high amounts of different methylotrophic organisms such as members of the genera *Methylothenera*, *Methylophilus* and *Methylovorus* were reported in the same study. These microbes were suggested to be related to the methanogenesis or degradation of chitin like substances (diversified macrofauna of the cave) and reflect the possible importance of particular single-carbon compounds in the cave environment.

1.5.2 Microbes and metal deposits

It is not uncommon that carbonate rocks host inclusions of metals, iron being the most frequently occurring one. In caves it can often be found in oxidized form in deposits or sediments and in reduced form in seepage or groundwater. Different iron oxidizers related to genera such as *Leptothrix*, *Gallionella* and *Siderooxidans* were often detected in sediments (Peck 1986), water (Moore 1981), stream cobble biofilms and surface of speleothems - secondary mineral deposits formed in a cave (Kasama and Murakami 2001). The above mentioned genera are thought to induce iron mineralization at the contact point of circumneutral anaerobic seeps and springs with oxygen. Bacteria were also shown to accelerate precipitation rates by up to 4 orders of magnitude on ferrohydrite speleothems (Kasama and Murakami 2001). Although the induced biomineralization and involvement of bacteria in the formation of iron ores is still under discussion (Baskar et al 2012, Wu et al 2009), the cave environment seems to provide the best preserved environment for the identification of promising geological records.

So-called ferromanganese deposits occurring in range of colors and hosting relatively high amount of clays, Al oxides, and others minerals (Spilde et al 2005), represent a special issue in caves. These types of deposits have been described from Lechuguilla Cave (USA), Wind Cave (USA) and Ochtina Aragonite Cave (Slovakia). The diversity of microorganisms in such formations is very high; in some deposits communities were even dominated by *Archaea* (Northup et al 2003). Moreover, common 16S rRNA genotypes were detected in DNA from recent deposits and from well-conserved paleosamples (Barton and Northup 2007). Although several microorganisms identified from ferromanganese deposits are affiliated with iron or manganese oxidizers, the mechanism of deposits formation still remains unclear.

1.5.3 Microbes and precipitation of carbonates

The complexity of the calcium carbonate precipitation phenomenon in cave environments has been addressed in many studies over the last two decades (Barton and Northup 2007, Castanier et al 1999, Forti 2002). Although caves provide a favorable environment for chemical precipitation (Castanier et al 1999, Ehrlich 1998) the role of microorganisms in this process is nevertheless discussed. Biologically induced and controlled mineralization of carbonates has frequently been reported from caves (Fig.3) (Barton et al 2001, Boquet et al 1973, Cañaveras et al 2001, Douglas and Beveridge 1998). Different mechanisms promoting precipitation have been proposed, e.g., the removal of crystallization inhibitors by bacteria (Bosak and Newman 2005), a shifting of the pH around the microbial cell through autotrophic processes, or nitrogen release and fixation (Cacchio et al 2004, Castanier et al 1999, Hammes and Verstraete 2002). Several researchers attribute calcite precipitation by bacteria to the active pumping of Ca^{2+} ions from inside the cell into the extracellular medium, a mechanism to avoid toxic levels of calcium concentrations (Cacchio et al 2004, Cai and Lytton 2004). Banks et al (2010) demonstrated that a *ChaA* knockout mutant strain lost expression of the Ca^{2+} ion efflux protein and was not able to growth on

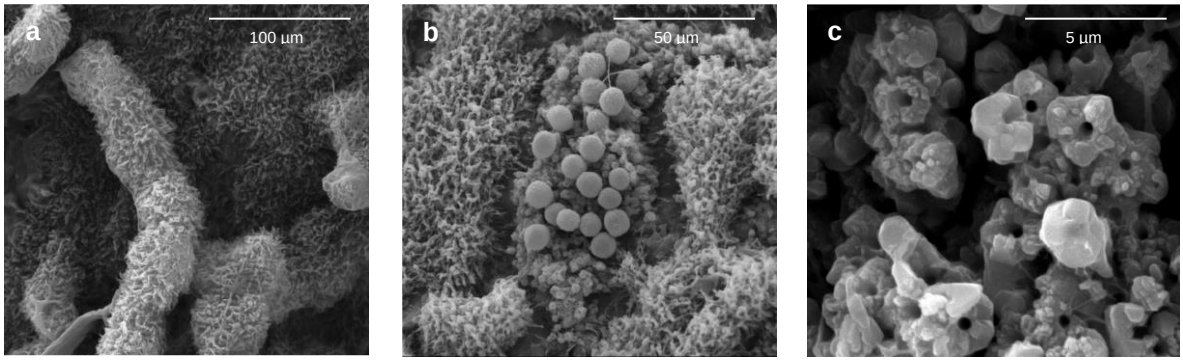


Figure 3: *Microbial communities associated with mineral precipitation in Altamira cave. Microphotographs of the microstructural organization (a) yellow colonies with $CaCO_3$ nest-like aggregates; (b) grey colony associated with spheroidal elements; (c) rosette-like aggregates (Cuezva et al 2009)*

carbonate rich media, as opposed to the wild type. Bacteria were proposed to play an important role for the initial crystal nucleation by these authors. This was further confirmed by Sanchez-Moral et al (2012), who suggested that bacterial activity would promote the first stages of deposition. Rusznyák et al (2012) showed taxon specific crystal formation in cultivation experiments with isolates obtained from speleothems of Herrenberg Cave (Germany). At the same time the progressive accumulation of carbonate was described as less favorable environment for bacterial growth (Sanchez-Moral et al 2012). The polysaccharides and different amino acids in the extracellular matrix of bacteria may also influence the crystal form, even if they are placed in abiotic carbonate rich environment (Braissant et al 2003). Altogether, microbial activity appears to be more important for the initiation of crystals and their shapes than for ongoing precipitation.

1.5.4 Microbes in aquifers of conventional karst

Despite a recent increase in microbial karst research the number of studies about pristine karst aquifers not related to the sulfur cycle remains very low. Most information originates from spring or ground water related studies (Farnleitner et al 2005, Gounot 1994 , Pronk et al 2009, Wilhartitz et al 2009), and the most frequently investigated aspect in karst systems is the microbiological monitoring of water safety, including the abundances and spreading of pathogen contaminates (Butscher et al 2011, Pronk et al 2007, Reischer et al 2008). Non-contaminated karst water seems to harbor relatively low bacterial abundances: The total numbers of planktonic *Bacteria* reported from pristine karst aquifers vary between 10^2 and 10^5 cells ml^{-1} (Griebler and Lueders 2009).

Studies on spring water often report microbial characteristics in the context of hydrological and chemical parameters such as discharge dynamics, turbidity, TOC and nitrate concentration. This allows a better understanding of interconnections between these parameters. For example, a correlation between turbidity of spring water, total bacteria counts, number of fecal contaminants, TOC and nitrate concentration was observed in karst outlets with fast response to flood events (Farnleitner et al 2005, Pronk et al 2006, Pronk et al 2009, Wilhartitz et al 2009). Nevertheless, a high temporal stability of microbial communities was also detected in karst springs, as determined by denaturing gradient gel electrophoresis (DGGE) fingerprints of 16S rRNA genes over a whole annual cycle (Farnleitner et al 2005). This has led to the proposal that there is an “autochthonous microbial endokarst community” (AMEC). Microbial habitat specificity has also been shown by Pronk et al., (2009) in a study about two related karst springs harboring water of three different origins (cold and warm karst ground water and percolating water from a swallow hole). The contribution of swallow hole water to spring discharge was demonstrated to be the most important factor for changes in the microbial community composition, and was

suggested as a provider of allochthonous microorganisms in the karst system. At the same time the possible surface origin of at least part of AMEC or the existence of metabolic active endokarst bacteria, e.g., in biofilms, remains under discussion. Data existing on microbial diversity in karst water is very scant and controversies. Several studies of microbial consortia composition of karst springs reported the dominance of α-,δ-, β-, γ-protobacterial sequences (Farnleitner et al 2005, Pronk et al 2009, Zeng et al 2010), but some researchers restrict AEMC to particular groups of δ-*Proteobacteria* and to less frequently detected *Nitrospira and Acidobacteria* due to the possible allochthonous origin of other bacterial phyla such as β- and γ-*Protobacteria* (Engel 2010, Pronk et al., 2009). At the same time it is doubtful if the chemolitoautotrophic metabolism associated with δ-*Proteobacteria* and *Nitrospira* linages is of importance in conventional karst aquifers, which are believed to depend on allochthonous organic matter (Culver 1985, Simon et al 2010). In general, planktonic microbes observed in springs were proposed to be dormant or inactive based on very low assimilation rates reflecting a possible influence of shearing forces (Farnleitner et al 2005, Pronk et al 2009, Wilhartitz et al 2009). By contrast, the relatively high heterotrophic activity in sediments of karst outlets indicates the possible dominance of biofilm-associated bacteria in such habitats. It has thus been suggested that the surface-attached fraction of the microbial communities plays a particularly important role in degrading allochthonous DOC and improving spring water quality. At the same time, autochthonus carbon sources, in particulary DOC released by viral lysis might also be a part of the trophic net in such systems (Wilhartitz et al 2009). Unfortunately there are currently no data about the transformations of aquatic microbial assemblages during transit through the karst systems, and it is difficult to estimate if the planktonic bacteria are active at particular conditions inside vadose endokarst water bodies. In summary, while at least some information is available about the phreatic part of karst systems (Griebler and Lueders 2009), the epiphreatic aquatic realms remain almost completely unexplored by microbiologists.

1.6 Karst research in Switzerland

The Institute of Swiss Speleology and Karstology (ISSKA) roughly estimates a Swiss reserve of karstic ground water of about 120 km^3. Such a volume corresponds to the combined volume of all Swiss lakes and forms 50% of all Swiss groundwater resources (ISSKA 2008).

This underlines the importance of karst research and favors the formation of research groups that are leading in the field. For instance, the collaboration between hydrologists and microbiologists at the University of Neuchatel has provided the first insight into the

microbiology of conventional karst aquifer (Pronk et al 2006, Pronk et al 2007, Pronk et al 2009).

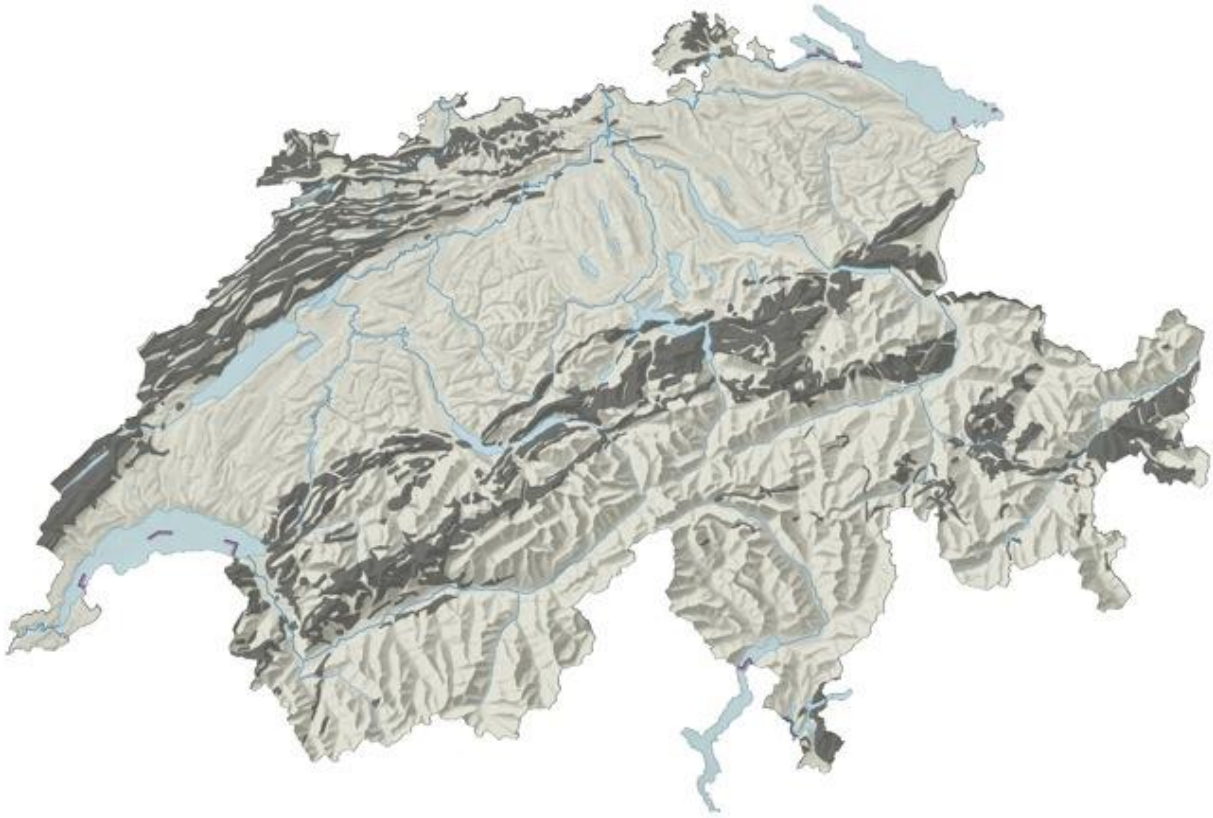


Figure 4:*Karst distribution in Switzerland (Atlas der Schweiz 2.0, 2004)*

Several projects on modeling of karst aquifers are currently active in some Swiss research institutions, e.g., SWISSKARST (http://www.swisskarst.ch) of ISSKA. Their aims include the precise estimation and prediction of karst water systems and their behavior by means of high-resolution models that are based on various parameters such as geological characteristics, porosity, precipitations, glacier melting, surface runoff etc. Data obtained from the above described project should eventually allow the proper simulation of spring discharge, flood prediction and the construction of 3D maps for karst hydrological systems, which are the prerequisites for suitable karst management.

Besides scientific research at academic institutions, a well-organised and highly motivated speleological society is active in Switzerland (Schweizer Gesellschaft für Höhlenforschung, SGH). SGH numbers approximately 1000 members and is indispensable in projects such as cave mapping, the creation of karst cadasters and collection of samples for scientific projects.

1.7 Study site

In this study, the majority of sampling and observations were conducted in Bärenschacht cave of Berner Oberland (Canton of Berne, Switzerland). It is located on the northern shore of Lake Thun close to the village of Beatenberg. The entrance was discovered in 1965 by two inhabitants of Beatenberg, and the cave was named Bärenschacht due to bear-bones founded in the first pit. The vertical part of the cave system was explored in 70ths to the sump at -565 m, which was passed by diving only in 1986. An unexpected amount of new labirintic passages was found beyond it. To allow the access to the new cave part for non-diving explorers, an artificial tunnel was constructed within 3 years (completed in 1995). The currently known maximal depth of Bärenschacht is approximately 970 m, and the cave counts more than 73 km of well-documented passages (August 2012, R. Siegenthaler). Although water income from the surface into the labyrinth is very scant, some sumps were identified in the lower part of the cave (Fig. 5). They are connected to phreatic streams, which seem to represent a hydrological link between large cave system Siebenhengste, located to the north-east from Bärenschacht, and springs of Lake Thun (Häuselmann 2002). The catchment area of this karst aquifers is shared between the Cantons Berne and Lucerne; it includes, amongst others, the Siebenhengste, Innerbergli and Schrattenfluh karren fields, forested areas and peat bogs. The rain events and snow melting in this region cause floods in the lower part of Bärenschacht. Highest water table fluctuations (up to 100 m) were observed in the North part of the cave. This epiphreatic zone is determined by flooding activity of the North Sump and hosts plenty of water pools with different flooding frequency. At the same time, hydrologically independent ponds fed by seepage water and characterized by distinct chemistry exist in the upper (vadose) part of the labyrinth.

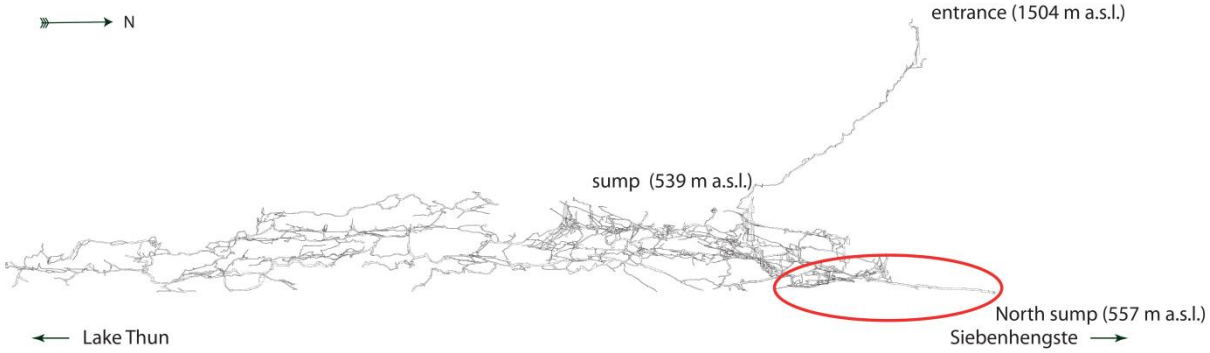


Figure 5: *Profile of the Bärenschacht cave. The epiphreatic zone affected by North sump is depicted in red. (Based on a survey by the Caving Association of the Hohgant Region (HRH), 2009)*

1.8 Aims of the study

The main goal of this study was to provide first insight into the microbiology of water in conventional endokarst. From the various karst water habitats, such as vadose and phreatic streams and different temporary water bodies, pools in the vadose and epiphreatic realms were chosen as study sites, since they represented accessible and more or less stable systems.

The first aim was to characterize the differences between ponds with diverse hydrology (epiphreatic vs. vadose). We planned to compare chemical and microbiological parameters between pools from zones with different water saturation and pools within the same zone, to assess the level of the microbial diversity and to define key players in such environments, in order to estimate the influence of the water chemistry on the microbial community composition, and to find possible interconnections with related habitats. In the context of these investigations, it was observed that the parameters of epiphreatic pools changed due to flood events and during the interjacent periods of stability. Therefore the subseqeunt project was to analyze post-flood dynamics and to determine the processes involved in possible community transformation. In particular, we addressed questions concerning the nature of floodwater inocula, the reproducibility of trends observed after floodings, and possible microbial transformation of DOC. Particular attention was paid to changes in diversity and community structure: We specifically tested the hypothesis that (i) bacterial community of flooding streams were mainly formed by stochastic processes, as previously proposed (Crump et al 2007, Holyoak et al 2005, Logue and Lindström 2008), and that (ii) the inocula mostly preshaped by mass effects were then subjected to the selective processes in stagnant epiphretic pools. Next Generation Sequencing was selected for this purpose in order to obtain enough statistical power to detect the concomitant phylogenetic restructuring of the microbial assemblages,

2. Microbial diversity in subsurface karst pools (vadose vs. epiphreatic)

Manuscript I:

Karst pools in subsurface environments: collectors of microbial diversity or temporary residence between habitat types.

Environmental Microbiology, 2010, 12:1061-1074

Environmental Microbiology (2010) **12**(4), 1061–1074 doi:10.1111/j.1462-2920.2009.02151.x

Karst pools in subsurface environments: collectors of microbial diversity or temporary residence between habitat types

Tanja Shabarova and Jakob Pernthaler*
Limnological Station, Institute of Plant Biology, University of Zurich, Seestr. 187, CH-8802, Kilchberg, Switzerland.

Summary

We studied bacterial diversity and community composition in three shallow pools of a Swiss karst cave system with contrasting hydrological and hydrochemical properties. The microbial assemblages in the pools were remarkably different, and only one operational taxonomic unit of 16S rRNA genes (OTU, 97% similarity) was shared between the three of them (total OTU number in all pools: 150). Unexpectedly high microbial phylotype richness was found even in the two pools without groundwater contact and with low concentrations of organic carbon and total cell numbers (< 10^4 ml^{-1}). One of these seepage water fed systems harboured 15 distinct OTUs from several deeply branching lineages of the candidate phylum OP3, whereas representatives of this group were not detected in the other two pools. A tentative phylogeographic analysis of available OP3-related sequences in the context of our data set revealed that there was generally little agreement between the habitats of origin of closely related sequence types. Two bacterial clades affiliated with the obligate methylamine utilizer *Methylotenera mobilis* were only found in the pool that was exposed to repeated flooding events. These bacteria formed relatively stable populations of up to 6% of total cell counts over periods of several months irrespective of inundation by groundwater. This suggests that karst water may provide a means of transport for these bacteria from terrestrial to freshwater habitats.

Introduction

The term 'karst' is used to describe a special type of landscape that developed due to the dissolution of bedrock, such as carbonate, marble and gypsum. The karstified regions are characterized by well-developed subterranean water drainage systems. Although large areas of the world are influenced by karst and important processes such as groundwater recourse formation depend on the karst environment (Manda and Gross, 2006; Ford and Williams, 2007), little is known about microbial community composition and heterotrophic processes in these habitats. While research on springs and boreholes can give us a first idea about the diversity and function of microbial assemblages in karst environments (Personne *et al.*, 2004; Pronk *et al.*, 2009), this is often insufficient for understanding microbial successions or transport in the water during its transformation from rain to ground or spring water. Since caves allow access to most types of natural subsurface formations they can be helpful for the study of undisturbed karst environments.

Cave pools represent interesting habitats to explore the diversity and strategies of bacteria in oligotrophic environments. Such habitats are typically low in dissolved organic matter (DOM) (Laiz *et al.*, 1999; Levy, 2007; Simon *et al.*, 2007), and they are without phototrophic eukaryotic production, which supplies an important proportion of the readily available DOM in surface waters (Hama and Yanagi, 2001). Heterotrophic microorganisms that thrive in fully oxygenated subsurface pools thus probably depend on other sources of OM to meet their carbon demand, e.g. compounds with no C–C bonds (methylotrophy), or recalcitrant humic substances from terrestrial sources (Einsiedl *et al.*, 2007).

Another interesting aspect of studying microbes in cave pools lies in the potentially large physicochemical and hydrological differences between closely neighbouring systems. Pools that are located within the so-called epiphreatic zone of cave systems are exposed to irregular flooding events by the rising groundwater table at a scale of weeks to months. Since the development of a specific 'local' bacterioplankton flora in standing water bodies is only observed if water residence times are relatively long (> 200 days) (Lindström *et al.*, 2006), the bacterial assemblages of such pools most likely closely resemble the microbial communities that are transported by the karst water stream. In contrast, pools in the so-called vadose zone are hydrologically much more confined systems that

Received 24 September, 2009; accepted 1 December, 2009. *For correspondence. E-mail pernthaler@limnol.uzh.ch; Tel. (+41) 44 716 12110; Fax (+41) 44 716 1211.

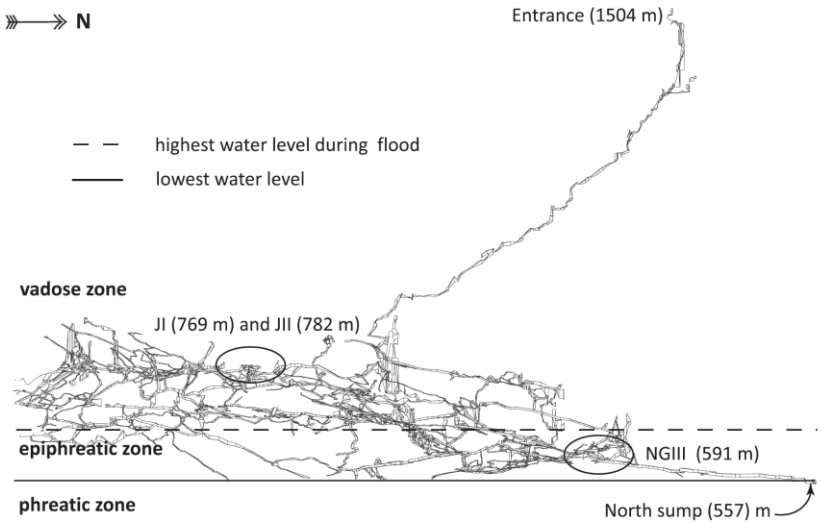


Fig. 1. Profile of the Bärenschacht cave system (northern part) and location of the studied pools NGIII, JI and JII. The broken line indicates the separation between the regularly flooded (epiphreatic) zone and the permanently dry (vadose) realms of the cave. Lake Thun is located to the south of the cave. Based on a survey by the Caving Association of the Hohgant Region HRH, 2009.

are exclusively fed by dripping or seeping water from rock surfaces, and only a very restricted exchange of microbial biota between such habitats can be envisaged. Moreover, these pools may substantially vary in water chemistry due to small-scale differences in geologic conditions, which in turn may affect microbial community composition (Barton *et al*., 2007).

The Bärenschacht is a large 950-m-deep cave system with 67 km of well-documented passages situated in the Bernese Oberland on the north side of Lake Thun (Switzerland). The multiple, vertical entrance pits and steeply dipping galleries cross flysch and sandstone layers before they lead into a limestone labyrinth (Fig. 1). Since surface rock formations are resistant to karstification and the water flow from surface is scant, bottom aquifers of Bärenschacht are almost exclusively fed by water from the neighbouring cave systems Siebenhengste and Schrattenfluh, situated in the northern part of the same catchment area. Thus, Bärenschacht forms the hydrological link between these systems and the underwater springs of Lake Thun (Häuselmann, 2002). The cave features eight routes to the groundwater table. The largest one is exposed to frequent and pronounced flooding, resulting in rises of the water table by up to 100 m. For most of the numerous water ponds located in the epiphreatic zone, this flood water is the sole feeding source. At the same time, several small ponds located in the highest (vadose) galleries of labyrinth are only supplied by seepage water.

We studied the diversity of bacteria in water of three pools from the epiphreatic and vadose zone of the Bärenschacht cave system in the context of hydrological and hydrochemical properties. Specifically, we wanted to assess if most of the taxa were shared between the systems or if the pools each featured unique microbial assemblages. In addition, we repeatedly quantified the community contribution of a phylogenetic lineage of methylotrophic bacteria in one of the pools over a period of 3 months.

Results

Physicochemical parameters

A detailed description of the three studied cave pools is given in *Experimental procedures*. The temperature in the pools was almost constant and ranged between 6.4°C and 6.8 °C (data not shown), and pH remained circumneutral in all three systems (Fig. 2). Four flooding events were recorded during the study period in the epiphreatic pool NGIII. The maximum levels of the water table exceeded the bottom level of NGIII by 3 m (1–3 March 2008) and 16 m (5–6 October 2007), thus completely exchanging the volume of the approximately 0.5-m-deep pool. In general, NGIII showed a higher variability of all measured chemical parameters (Fig. 2 and Table 1). Total organic carbon (TOC) in NGIII ranged between 1.9 and 4.5 mg l^{-1} and was up to 12 times higher than in the vadose pools JI and JII. In contrast, conductivity as well as the concentrations of all observed ionic compounds and of silicate were considerably higher in JI and JII than in the epiphreatic NGIII system. The calculated saturation index for calcite was positive for pools JI and JII (between 0.3 and 0.4), indicating that precipitation was still ongoing in these systems, whereas it showed negative values for NGIII (between –0.3 and –0.1). Although the conductivity of both vadose pools was in the same range, the ionic composition of water from these sites was considerably different. Concentrations of nitrate and calcium were higher in JI (~2 and ~1.3 times respectively), and magnesium and sulfate concentration were higher in JII (~4.2

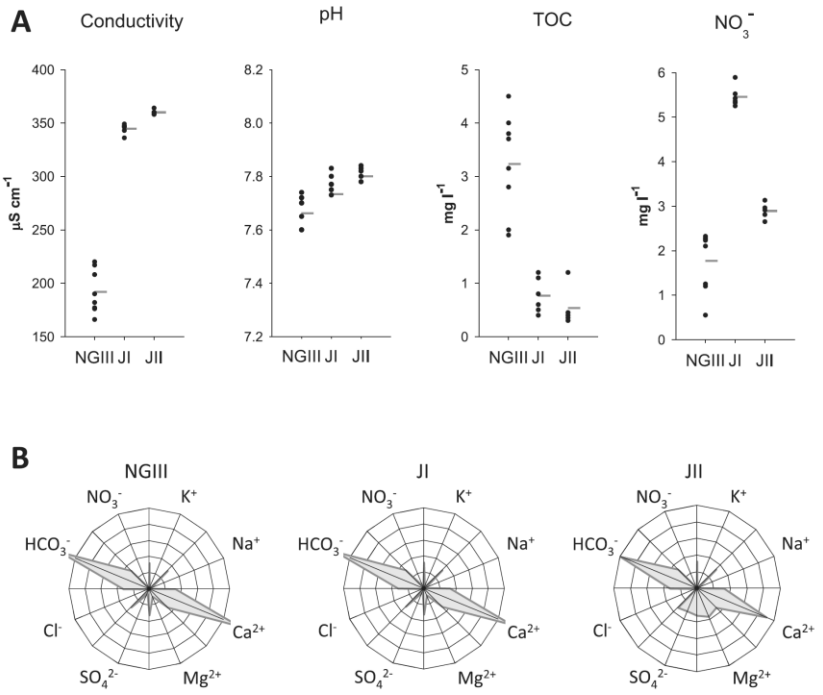


Fig. 2. Selected physicochemical characteristics of the pools (see also Table 1).
A. Ranges of conductivity, pH and concentrations of TOC and nitrate (symbols) and average values (dashed line) of five to eight measurements.
B. Maucha ionic diagrams presenting the ratios of the equivalents of the different ions (shape of shaded area). The origin of the diagram depicts ionic isoequivalence. The shaded area is of equal size in all diagrams and represents 100% of the equivalents of all depicted ions.

and ~5.2 times respectively). The ion ratios in NGIII and JI were nearly identical, but differed significantly from those in JII, where gypsum also appeared to play an important role in affecting water chemistry (Fig. 2B) (Häuselmann, 2002).

Diversity analysis and phylogenetic affiliation of 16S rRNA genes

Altogether 339 16S rRNA gene sequences were obtained, 22 of which were discarded as chimeras, five as plant chloroplasts (in pool NGIII, possibly introduced during flooding) and five as mitochondria. In addition, four sequences were excluded that were most closely related to known contaminants (Tanner *et al.*, 1998; Barton *et al.*, 2006). The final collection consisted of 100 almost complete sequences (> 1400 nucleotides) from NGIII, 100 from JI and 103 from JII.

Table 1. Hydrochemical parameters in the three pools.

	JI	JII	NGIII
SO_4^{2-} (mg l^{-1})	8.6 (0.6)	45 (1)	6.25 (1.5)
Cl^- (mg l^{-1})	1.6 (0.1)	1 (0.1)	< 0.5
Alk (meq l^{-1})	3.6 (0.1)	3.1 (0.1)	2.0 (0.3)
Na^+ (mg l^{-1})	< 2.5	3.4 (0.1)	< 2.5
K^+ (mg l^{-1})	2.8 (0.1)	< 1	< 1
Ca^{2+} (mg l^{-1})	72 (1)	57 (1)	41 (4)
Mg^{2+} (mg l^{-1})	3.3 (0.17)	14.0 (0.15)	< 2.5
H_2SiO_4 (mg l^{-1})	8.8 (0.4)	9.9 (0.1)	3.1 (0.5)

Values are averages of two to four separate samplings, and standard deviations are given in brackets. Alk, Alkalinity.

The apparent community overlap between the three closely neighbouring systems was conspicuously low. Only one shared operational taxonomic unit (OTU, similarity level: 97%) could be detected in all pools (Fig. 3A); it was affiliated with a subgroup of the genus *Rhodoferax*. More specifically, a common monophyletic group within this OTU was formed by nine sequences from NGIII and seven from JI, whereas the two members of this OTU originating from JII were phylogenetically separated (as suggested by high bootstrap values, data not shown). One additional OTU affiliated with the genus *Flavobacterium* was shared between pools NGIII and JII. The four OTUs shared between pools JI and JII were affiliated with *Oxalobacteraceae* (two OTUs), with *Acinetobacter* (one OTU) and with *Nitrosomonadaceae* (one OTU) (Table 2). In addition to assessing OTU overlap, θ values were calculated to compare the similarity of communities based on the relative abundances of sequences within OTUs (Yue and Clayton, 2005). This analysis indicated that despite the different numbers of shared OTUs, the similarity of the communities in pools JI and JII at the 97% level was in the same range as in JI and NGIII. In contrast, the community similarity between pools JII and NGIII was much lower (as reflected by a lower θ value) and corresponded well with the low numbers of shared OTUs (Fig. 3A).

High total and local phylotype richness was found in the studied systems, as reflected by the Chao 1 estimator

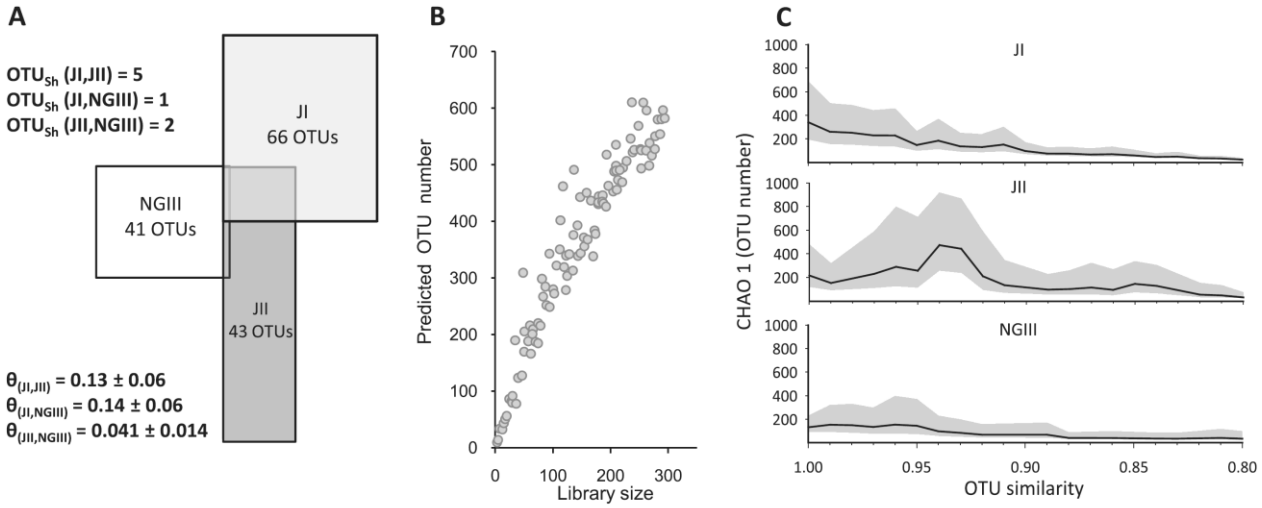


Fig. 3. Diversity and similarity of microbial 16S rRNA genes from the studied systems.
A. Overlap of the numbers of OTUs in the systems and pairwise community similarity indices (θ values) calculated for a sequence similarity level of 97%.
B. Relationship of total OTU richness estimates (Chao 1, 97% sequence similarity) with subsample size of the pooled sequence collection from all three ponds.
C. Estimate of the total OTU richness (Chao 1) at different levels of sequence similarity in each of three libraries.
OTU$_{Sh}$: numbers of OTUs shared between two systems.

(Fig. 3B and C). A substantial undersampling of community diversity was moreover suggested by the non-asymptotic progression of Chao 1 values in randomly drawn subsets of the libraries from the individual pools (data not shown) and of the combined sequence collection (Fig. 3B). Interestingly, the values of Chao 1 in pool JII displayed a clear maximum at a cut-off of 94% of sequences similarity (i.e. approximately below the genus level) (Fig. 3C), hinting at an unusual phylogenetic structuring of the diversity in this system. Indeed, this maximum was not discernible if sequences affiliated with β-*Proteobacteria* were excluded from the analysis (data not shown).

The phylogenetic affiliation of sequences of all three systems is summarized in Table 2. Besides β-*Proteobacteria* (17 clones), a high number of sequence types in pool JI were affiliated with *Actinobacteria* (14 clones) and with candidate phylum OP3 (16 clones). Genotypes from that latter lineage were exclusively present in JI and were affiliated with at least three deeply branching subclusters (Fig. 4). Although β-*Proteobacteria* were the most abundant set of sequence types in NGIII and JII, nearly no phylogenetic overlap within this cluster between pools was observed: members of the family *Methylophilaceae* were exclusively found in NGIII (21 clones). A more detailed examination of these sequence types revealed the presence of bacteria from two clearly separated phylogenetic lineages (*Methylophilaceae* I and *Methylophilaceae* II) which were both related to *Methylotenera mobilis* (< 4% of sequence difference) (Fig. 5). In contrast, sequences affiliated with *Oxalobacteraceae* family were found only in the two crystalline pools JI and JII and clearly dominated the library from the JII system (40 clones).

Microscopic analysis

No quantitative microscopic analysis of bacteria on membrane filters was possible in samples from JI and JII, because (i) microbial cell numbers in these two pools were extremely low (< 10^4 cells ml^{-1}) and (ii) the water from these systems contained high numbers of inorganic particles that impeded the evaluation of more concentrated samples.

Bacterial abundances in eight samples from pool NGIII ranged between 2.7×10^5 ml^{-1} and 5.2×10^5 ml^{-1}. The detection rates of bacteria by fluorescence *in situ* hybridization and catalysed reporter deposition (CARD-FISH) varied from only 33% of total cells stained with 4′-6-Diamidino-2-phenylindole (DAPI) in November 2007 to 65% in January 2008. Only between 49% and 75% of all hybridized *Bacteria* (14–24% of DAPI counts) could be further identified by the applied group-specific probes. β-*Proteobacteria* represented the most abundant bacterial lineage that could be identified by FISH in all samples (mean 48 ± 7% of *Bacteria*, standard deviation ± 1), followed by microbes affiliated with the *Cytophaga–Flavobacteria* cluster (8 ± 2% of *Bacteria*) (Fig. 6A). The discrepancy between the frequency of these bacteria in the clone libraries and their apparent rarity in the

Table 2. Phylogenetic affiliation of bacterial 16S rRNA genes obtained from the studied pools.

Division and order	Family	Genus	NGIII	JI	JII
Acidobacteria			1	2	1
Actinobacteria			1	14	10
Actinomycetales	*Microbacteriaceae*	*Cryobacterium*	1		
	Intrasporangiaceae	*Janibacter*		3	
	Frankiaceae	*Frankia*			1
	Sporichthyaceae	*Sporichthya*			8
Micrococcales	*Micrococcaceae*	*Arthrobacter*		5	
Rubrobacterales	*Rubrobacteraceae*			6	
Unclassified					1
Bacteroidetes			23	4	8
Flavobacteriales	*Flavobacteriaceae*	*Flavobacterium* (1)	10		8
		Chryseobacterium		4	
	Cytophagaceae	*Flexibacter*	5		
		Flectobacillus	5		
	Chitinophagaceae	*Chitinophaga*	2		
	Unclassified		1		
Firmicutes				2	
Fibrobacteres			2		
Nitrospirae	*Nitrospiraceae*	*Nitrospira*		7	1
Proteobacteria					
α-Proteobacteria			1	3	7
Rhodospirillales	*Rhodospirillaceae*		1		5
Rhizobiales	*Brucellaceae*			1	
	Hyphomicrobiaceae	*Hyphomicrobium*		1	1
		Devosia			1
	Xanthobacteraceae			1	
β-Proteobacteria			39	17	49
Burkholderiales	*Comamonadaceae*	*Aquaspirillum*	1		1
		Acidovorax		1	1
		Rhodoferax (1)	14	8	3
		Polaromonas			1
	Oxalobacteraceae (2)			7	40
Methylophilales	*Methylophilaceae*	*Methylotenera*	21		
Nitrosomonadales	*Nitrosomonadaceae* (1)			1	1
Rhodocyclales	*Rhodocyclaceae*	*Sterolibacterium*			1
Unclassified			3		1
γ-Proteobacteria			13	5	7
Oceanospirillales			8		
Alteromonadales		*Teredinibacter*			1
Pseudomonadales	*Pseudomonadaceae*	*Cellvibrio*			1
	Moraxellaceae	*Acinetobacter* (1)		2	3
Xanthomonadales	*Coxiellaceae*	*Aquicella*	2		
		Coxiella			1
		Unclassified		1	
	Sinobacteraceae	*Nevskia*		2	1
Methylococcales	*Methylococcaceae*	*Methylobacter*	3		
δ-Proteobacteria			4	5	4
Desulfobacterales	*Nitrospinaceae*			1	2
Myxococcales	*Nannocystaceae*	*Nannocystis*		1	
Syntrophobacterales	*Syntrophaceae*	*Syntrophus*	4		
Unclassified				3	2
BRC1				2	
OD1			2	1	
OP3				16	
OP5				3	1
OP10					6
OP11				3	
Planctomycetes			2	2	1
TM7			6	7	3
Verrucomicrobia			3	1	1
Unclassified			3	6	3

The numbers of OTUs (97% similarity) within a phylogenetic group that are shared between two pools are shown in parenthesis. Sequence types were assigned to the lowest possible taxonomic level by phylogenetic analyses (not shown) and taxonomic categories were used from the List of Prokaryotic Names with Standing in Nomenclature (http://www.bacterio.cict.fr/).

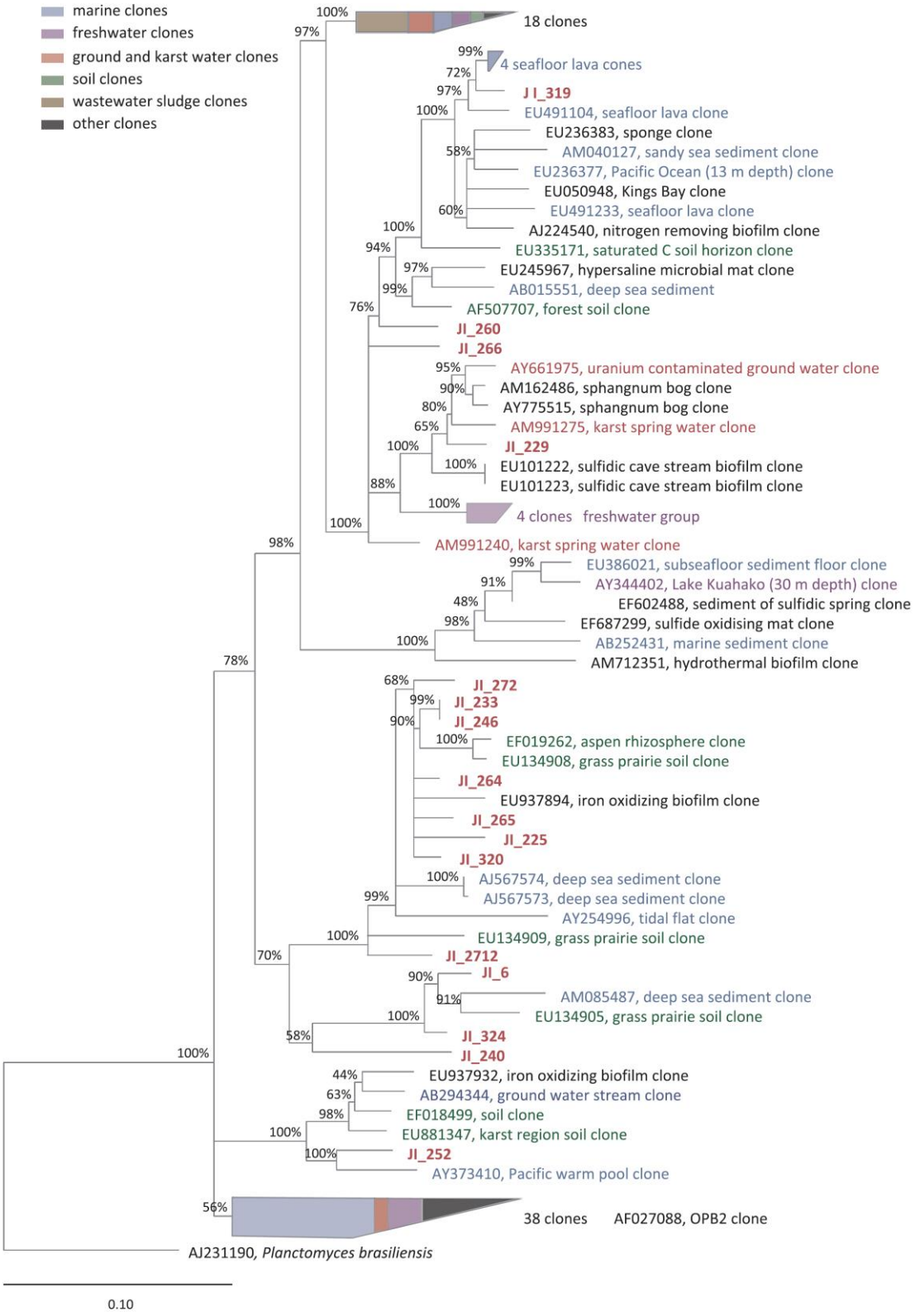


Fig. 4. Phylogenetic affiliation of members of the candidate phylum OP3 from pool JI. Sequences from public databases were colour coded according to their origin. Nodes with values < 50% of bootstrap support (100 maximum likelihood trees) were collapsed into multifurcations.

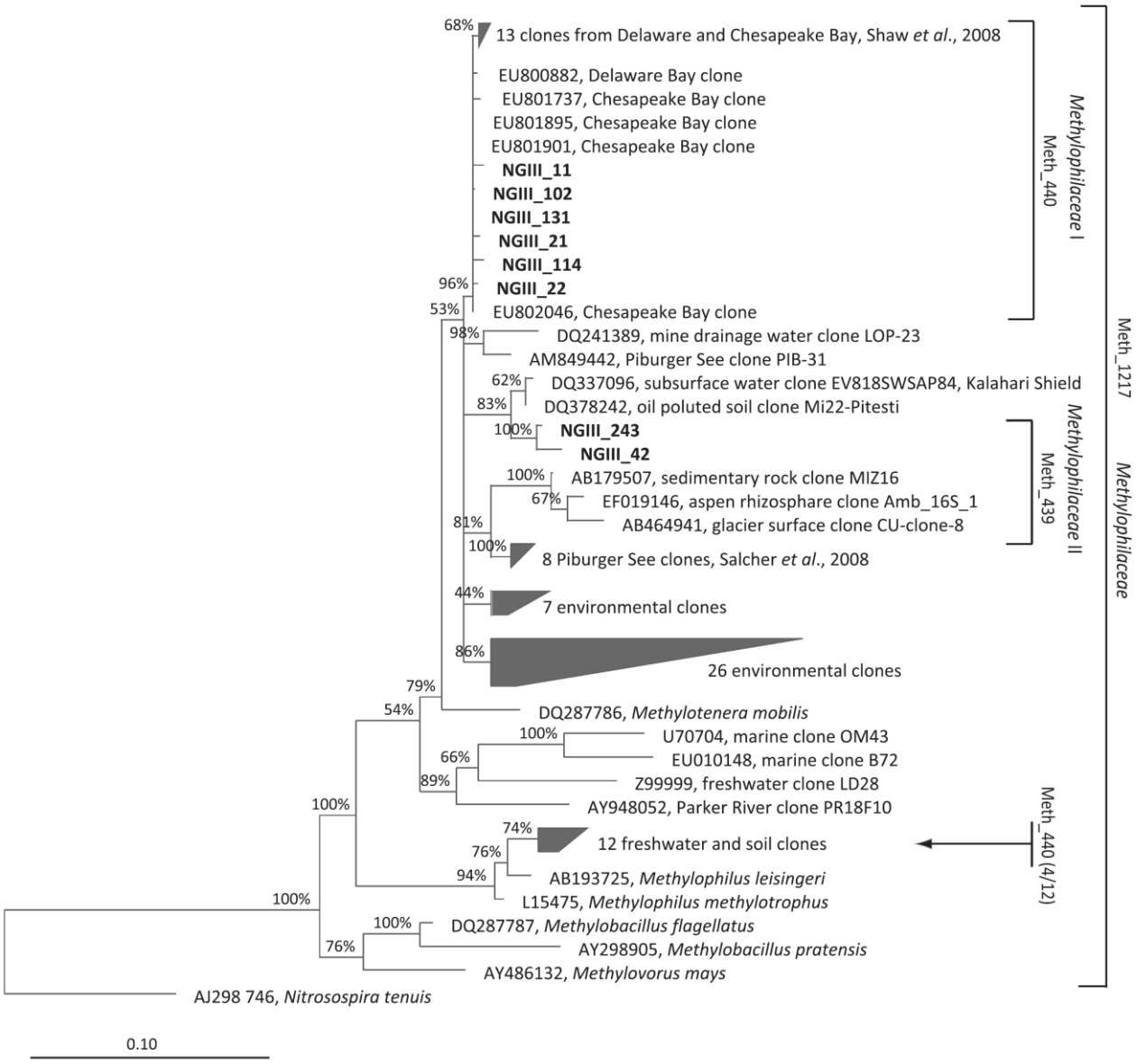


Fig. 5. Phylogenetic affiliation of sequences affiliated with *Methylophilaceae* obtained from pool NGIII. The sequences targeted by probe Meth_1217 and by two newly developed probes for different populations from the *Methylotenera mobilis* clade are marked by brackets. Nodes with values < 50% of bootstrap support (100 maximum likelihood trees) were collapsed into multifurcations.

environment might in parts be due to insufficient coverage of the *Cytophaga–Flavobacteria* lineage by the CF319a probe (Amann and Fuchs, 2008). The proportions of α-*Proteobacteria* never exceeded 6% of hybridized cells (4% of DAPI signals), γ-*Proteobacteria* could only be found in three of eight samples and *Actinobacteria* only in two (data not shown).

Using a specific probe for members of the *Methylophilaceae* family (Met 1217) (Friedrich *et al.*, 2003) up to 13% of *Bacteria* (11.8 ± 0.8%) could be detected in the five samples from pool NGIII that were evaluated with this probe (Fig. 6B). Consequently, two FISH probes were designed to determine the proportions of cells from both groups of *Methylophilaceae* from the *M. mobilis* lineage (Fig. 5) in NGIII. Since it was not possible to create a probe that would only target the newly obtained sequences of the *Methylophilaceae* II clade, the probe Meth_439 was designed to also detect bacteria from another phylogenetically closely related group (as depicted in Fig. 5) (probe sequence: 5′-GAC TAA GTG TTC TTC CCT TGC G-3′, competitor sequence 5′-GAC TAA GTT TTC TTC CCT TGC G-3′, optimal formamide concentration in hybridization buffer: 55%). The second probe, Meth_440, was originally designed to exclusively

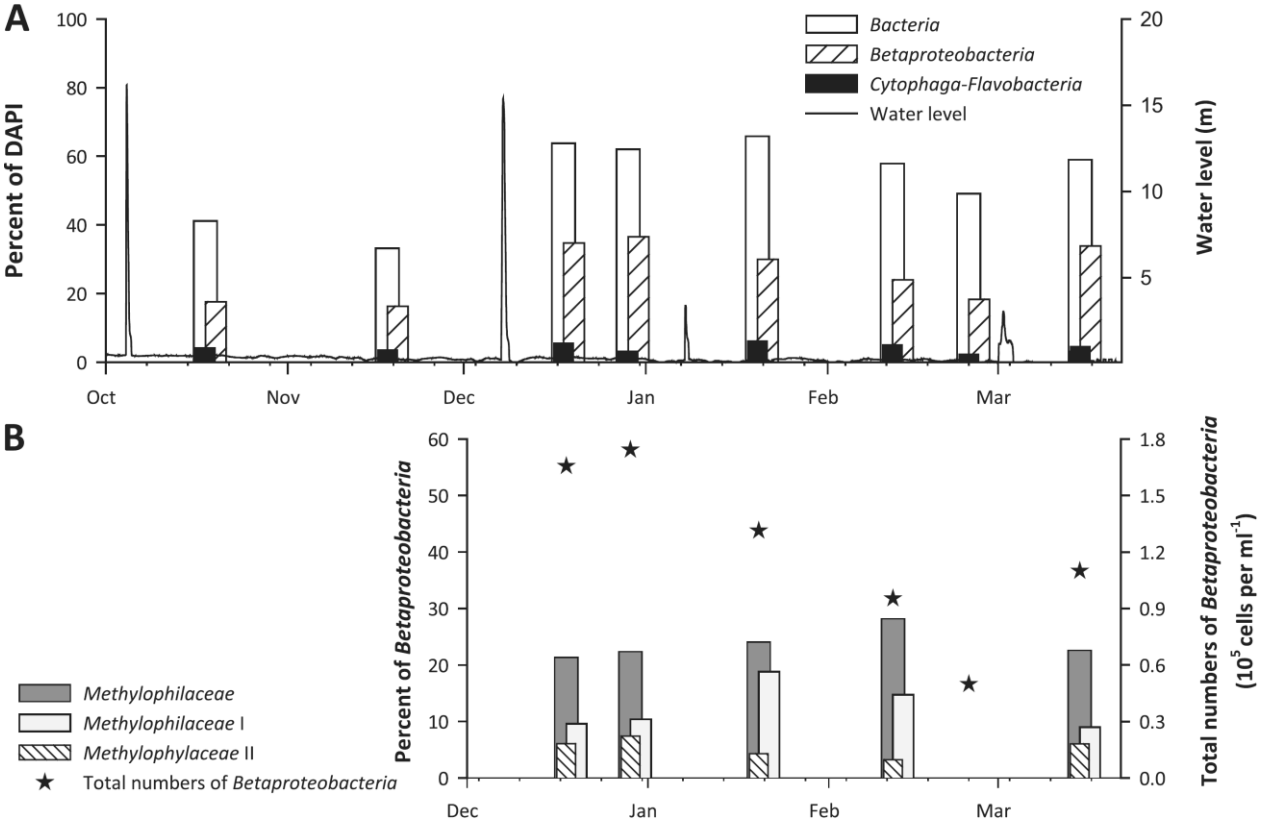


Fig. 6. Changes of water level and of bacterial composition in the epiphreatic pool NGIII.
A. Proportions of cells detected with FISH probes for large phylogenetic groups.
B. Proportions of cells detected with probes for *Methylophilaceae* and for two subpopulations related to *Methylotenera mobilis*, and total numbers of cells detected with a probe for β-*Proteobacteria*.

target the much larger *Methylophilaceae* I clade (probe sequence: 5′-AGA AAG TAA GTT TTC TTC CCT TGC-3′, competitor sequence: 5′-AGA ATG TAA GTT TTC TTC CCT TGC-3′, optimal formamide concentration in hybridization buffer: 55%). After completion of the study, a re-testing of probe specificity on a more recent version of the SILVA database (rel. 98) (Pruesse *et al.*, 2007) showed that the probe also binds to four sequences affiliated with close relatives of *Methylophilus leisengeri* (Fig. 5). Nevertheless, both *in silico* and *in situ* testing showed that the two probes can readily discriminate between the newly obtained sequences from the *Methylophilaceae* I and II clades respectively. They were thus suitable to prove the co-occurrence of the two groups in pool NGIII.

Bacteria affiliated with both the *Methylophilaceae* I and II lineages could be observed in NGIII throughout the study period (Fig. 6B). The two populations significantly differed in their relative abundances if regarded over the entire study period (*t*-test: $P < 0.05$), but this difference was within the typical error range of the quantification protocol at some time points (e.g. 29 December 2007). The total proportions of bacteria from both subclusters accounted for 63% (12 February 2008) to 95% (20 January 2008) of all *Methylophilaceae* detected by probe Met 1217. Although between 33% and 65% of β-*Proteobacteria* were found to incorporate radiolabelled thymidine, as assessed by microautoradiography (MAR), no uptake of this tracer by members of the *Methylophilaceae* family could be detected at any time point (data not shown).

Discussion

Diversity analysis

The three study systems featured high richness both of unique phylotypes and of OTU with similarity levels ranging from 99% to 93% (Fig. 3C). Each individual library and the pooled set of sequences from all three systems proved to be insufficiently large to yield precise diversity estimates (Fig. 3B) (Kemp and Aller, 2004). This is surprising, since estimates of phylotype richness in extreme environments or in surface water bacterioplankton can

typically be obtained in collections of < 100 rRNA gene sequences (Kemp and Aller, 2004). In contrast, an undersampling as pronounced as our systems is a more characteristic feature of habitats with high microbial densities, e.g. wastewater bioreactors or marine sediments (Kemp and Aller, 2004). However, the cell numbers in the waters of the vadose pools JI and JII were at least five orders of magnitude lower than in such systems, indicating that most of the numerous co-occurring genotypes were present in very low densities.

The two pools thus appeared to function as reservoirs of microbial diversity rather than favouring the exclusive dominance of a few particularly well-adapted species, as observed in extreme environments (Tyson *et al*., 2004). At the same time, these genotype collections were highly characteristic for the respective system, as reflected by almost 90% unique OTUs in each pool, the equally low weighted pairwise similarity (θ values, Fig. 3A), or the specific occurrence of bacteria from the OP3 phylum candidate JI (Table 1).

Since JI and JII are in close vicinity (Fig. 1) and bacterial cells can be dispersed within cave systems by various mechanisms (Mulec, 2008), this contrasting microbial community composition might be related to the conspicuous differences in water chemistry (Fig. 2, Table 1). The high concentrations of SO_4^{2-} and Mg^{2+} in JII at comparable levels of conductivity in both pools point to the potential importance of local geochemistry for the presence or lack of particular bacterial groups in the water of cave pools, as has also been described for rock-associated microbial assemblages (Barton *et al*., 2007). However, it remains unclear if (and which of) the numerous taxa in the vadose pools were mostly dormant or metabolically active. It is possible that some of this apparent high diversity might have originated from dead bacteria or even from remnant dissolved DNA that is known to persist in some freshwater environments (Matsui *et al*., 2001). At the same time, it appears unlikely that substantial amounts of degradable dead cells would accumulate in the extremely TOC-poor waters of the cave pools (Fig. 2A), and most dissolved DNA probably passes through the membrane filters used for biomass concentration.

In contrast, the high diversity in the epiphreatic pool NGIII is best interpreted in the context of the repeated flooding events (Fig. 4), which probably collect bacteria from a variety of oxic or anoxic habitats such as soils (Balkwill *et al*., 1998). The microbial assemblages in NGIII were strikingly different from those in either of the vadose pools (Fig. 3A), suggesting that only very few microbial genotypes from the karst water of the Bärenschacht system were indeed able to colonize and/or subsist in the latter environments. The only OTU shared between all three systems was affiliated with the genus *Rhodoferax*. This phylogenetic group appears to thrive in a variety of subsurface habitats, and closely related sequence types have been retrieved both from groundwater (Griebler and Lueders, 2009; Pronk *et al*., 2009) and from cave rock surfaces (Barton *et al*., 2007). It should be noted that the low fraction of shared OTUs between the three systems may also in parts be a consequence of undersampling (Fig. 3B), and some common community members might have been missed due to their rarity. An exhaustive assessment of phylotype richness appears only feasible by other approaches, e.g. massively parallel tag sequencing (Sogin *et al*., 2006).

While our libraries were too small to allow precise richness estimates, it is nevertheless possible to confine the range within which these estimates fell and thus to reveal additional differences in the phylogenetic structuring of diversity between the study systems (Fig. 3C). Most strikingly, the microbial assemblages in JII featured a distinct maximum of predicted OTU numbers at a sequence similarity level of 94%, which was absent in both other pools, and led to a significantly higher OTU prediction in JII than in NGIII (i.e. non-overlapping 95% confidence intervals). A local maximum of OTU richness at 94% similarity was also observed using a different diversity measure (ACE, data not shown). While this unusual pattern might be only a computational artefact, it also seems to reflect that the sequence collection from JII indeed featured roughly equal shares of β-*Proteobacteria* affiliated with *Oxalobacteraceae* and of very deeply branching phylogenetic groups that were only represented by a single genotype. This may indicate that *Oxalobacteraceae* as a group were particularly favoured by the ambient growth conditions in the pool. While these bacteria are not regarded as typical inhabitants of groundwater (Griebler and Lueders, 2009), some *Oxalobacteraceae* were common in 16S rRNA gene clone libraries from rock surfaces (Barton *et al*., 2007) and sulphate-rich groundwater wells (Hwang *et al*., 2009).

Candidate phylum OP3

Representatives of the candidate phylum OP3 (Fig. 4) of the *Planctomycetes–Verrucomicrobia–Chlamydiae* superphylum (Wagner and Horn, 2006) were originally discovered in samples from an acidic hot spring (Hugenholtz *et al*., 1998), but have since then been reported from oxic, sub- and anoxic environments such as sediments (Musat *et al*., 2006), soils (Derakshani *et al*., 2001), biofilms (Egli *et al*., 2003), waste water and sludge (Shigematsu *et al*., 2006), surface freshwaters (Donachie *et al*., 2004), and ground- and karst waters (Griebler and Lueders, 2009; Pronk *et al*., 2009). So far, there are no isolates or metagenomic data of microbes related to OP3. Fifteen distinct OTUs from the vadose pool JI were affiliated with this group, and our data from a single system contribute > 10% to all currently available almost

complete OP3 rRNA gene sequences. This indicates that the diversity within this phylum might presently still be substantially underexplored.

While our phylogenetic reconstruction may be biased due to the limited data set, the high bootstrap support of most branchings in Fig. 4 nevertheless suggests that many monophyletic lineages within the OP3 candidate phylum seem to have radiated across a range of different environments. Almost no single clade was exclusively formed by sequences obtained from one habitat type, and there were conspicuous discrepancies in the origin of even closely related genotypes. For example, the OTUs from our study were most closely affiliated with phylotypes from the rhizosphere, from marine sediments or from acidified bog lakes, but less so to other sequences from subsurface environments. At the same time, pool JI harboured bacteria from several different deeply branching clades of the candidate phylum, despite the most likely rather restrictive environmental conditions in this environment. In view of these somewhat contradicting observations, it is intriguing to speculate if the OP3 candidate phylum might be but a vestige of an ancient phylogenetic group that once was substantially more diversified. Alternatively, it is also conceivable that the observed deep divergence might be due to evolutionary isolation or a specialization for particular microniches that cannot be revealed by our comparatively simple analysis of environmental parameters.

OP3 sequence types were only obtained from one of the vadose pools that featured significantly higher concentrations of NO_3^- (Table 1). Interestingly, the sequence collection from JI also yielded several phylotypes of nitrite-oxidizing bacteria affiliated with *Nitrospira* (Table 2). This co-occurrence has also been reported by several researchers studying ground and karst water (Alfreider *et al*., 2002; Pronk *et al*., 2009), sulfuric hot springs (Hugenholtz *et al*., 1998), cave soil (Zhou *et al*., 2007) or nitrogen-removing biofilms (Egli *et al*., 2003) and might thus represent a promising target for future research.

Methylophilaceae *in the frequently flooded cave pool*

Approximately 20–40% of β-*Proteobacteria* in NGIII were *Methylophilaceae* (Fig. 6) from at least two distinct lineages affiliated with *M. mobilis* (Kalyuzhnaya *et al*., 2006) (Fig. 5). Although sequence types affiliated with *Methylophilaceae* have been found in subsurface waters before (Tian *et al*., 2005), these bacteria are currently not regarded as typical groundwater inhabitants (Griebler and Lueders, 2009). The type strain of *M. mobilis* has been recently isolated from Lake Washington sediments (Kalyuzhnaya *et al*., 2006); it utilizes methylamine as a sole carbon source both under oxic (Kalyuzhnaya *et al*., 2006) and microaerophilic (Kalyuzhnaya *et al*., 2008) conditions. Large populations of *Methylophilaceae* and 16S rRNA gene sequences from the *M. mobilis* clade have been found in the suboxic and anoxic layers of a lake (Salcher *et al*., 2008). In contrast, other phylotypes affiliated with this species were present on the leave and root surfaces of aquatic angiosperms (Crump and Koch, 2008). These bacteria thus appear to be able to thrive over a large range of oxygen concentrations. Their activity might depend both on anaerobic processes such as methanogenesis (Chen *et al*., 2009) and on interactions with higher plants. In addition, it is conceivable that these species also play a role in the turnover of recalcitrant halogenated organic compounds (Fetzner and Lingens, 1994; Grimvall and de Leer, 1995), or in the utilization of products from polysaccharide degradation (Chen *et al*., 2009).

The subclade that harbours the majority of phylotypes from NGIII (*Methylophilaceae* I, Fig. 5) also contains a large number of closely related sequence types from the Chesapeake Bay (Shaw *et al*., 2008). Over 50% of the total volume of water in streams of the Bay watershed is of groundwater origin (Bachman *et al*., 1998). The bedrock in the Bay catchment area features high amounts of carbonates and predominantly consists of forested territories. This is similar to the geology and vegetation at our study site, and the Bärenschacht system forms a direct hydrological link between neighbouring cave systems in the same catchment area and the underwater springs of Lake Thun (Häuselmann, 2002). Repeated observations of plant material (fir needles and root parts) in the epiphreatic pool as well as the relative high TOC values in this habitat (Fig. 2A) also point to its close connection with terrestrial habitats. Direct microscopic analysis moreover indicated that the relative proportions of bacteria from the *Methylophilaceae* I clade in NGIII were virtually unchanged by the repeated flooding events (Fig. 6) in the absence of noticeable numbers of protistan grazers (T. Shabarova, unpubl. obs.) and that these bacteria did not incorporate radiolabelled thymidine. However, the latter observation might also indicate that the studied bacteria were not able to incorporate thymidine. Nevertheless, it appears likely that bacteria related to *M. mobilis* were probably passively transported into the pool by karst water influx, but did not substantially grow in this habitat over periods of several weeks.

In summary, our investigation shows that microbial assemblages in subsurface karst pools may be highly diverse. Pools with different hydrochemical properties also remarkably differed in community composition, hinting at pronounced habitat filtering. The investigation of a larger number of karst pools with both similar and contrasting hydrochemical and hydrological properties and high-throughput analyses tools, e.g. 454 tag sequencing (Rothberg and Leamon, 2008), would help to substantiate

this hypothesis. Moreover, depending on hydrological properties, karst pools appear to function either as static collectors of microbial diversity (vadose pools JI, JII) or as ephemeral residence (epiphreatic pool NGIII) during the transport of bacteria from terrestrial habitats to freshwaters or estuaries.

Experimental procedures

Sampling procedures

We have studied three pools located in the labyrinth part of Bärenschacht cave. The dimensions of these systems are similar: approximately 20 m long, 2 m wide and 0.5 m deep. The pool NGIII is situated in the limestone gallery of an epiphreatic zone of the cave system and the pools JI and JII are located in the vadose zone (Fig. 1). Both J pools contain secondary mineral deposits: JI in the form of rafts on water surface, and JII in the form of crystals on pool bottom. The pools can be reached from the surface within 4 h. Five hundred millilitres of water was collected monthly between October 2007 and March 2008 from all three pools for the determination of pH, conductivity, TOC and nitrate concentrations. Another 500 ml of water was obtained at bimonthly intervals to assess the concentration of sulfate, chloride and cations. All samples for physicochemical analyses were collected in glass bottles pre-washed with 2 M HCl. All samples were delivered in the lab within 24 h. Temperature and water level in NGIII pool were monitored with a data logger placed at the bottom of the pool over the whole sampling period. Temperature in JI and JII was measured during the sampling procedure.

Water samples for bacterial cell counts and identification were collected with sterile syringes from a depth of 5 cm at least monthly. Samples were fixed with formaldehyde solution (FA, 2% final concentration) at the site. Microbial biomass for DNA extraction was gathered on 29 December 2007 and 20 February 2008 for NGIII and on 25 February 2008 for JI and JII. For this purpose, unfixed pool water (1.5–2.5 l) was filtered onto polycarbonate filters (type GTTP, pore size 0.2 μm, diameter 47 mm, Millipore) directly at the site. For transport the filters were placed into sterile tubes containing lysis buffer from the PowerSoil DNA Isolation Kit (MOBIO laboratories).

At five time points separate samples from the NGIII pool were prepared for later MAR analysis: triplicate subsamples of pool water (10 ml) plus one pre-fixed control were amended with [^{3}H]-thymidine (Amersham; specific activity 75.0 Ci mmol^{-1}, final concentration, 20 nM), incubated for 2–3.5 h at the site and fixed with FA as described above.

Physicochemical parameters

Conductivity and pH were measured in the laboratory with an InoLab benchtop meter (WTW) and a pH-Vision 6071 microcomputer (Jenco Electronics) respectively. Thirty millilitres of water sample was frozen for the subsequent determination of TOC on a TOC-5000 analyser (Shimadzu). Nitrate concentrations in pool water were measured by the cadmium reduction method after filtration through membrane filters (type GTTP, Millipore) that had been pre-rinsed with 2 M HCl and deionized sterile water (Milli-Q, Millipore). The concentrations of sulfate, chloride and of cations were assessed by ion chromatography (Compact IC 761, Metrohm). The alkalinity was measured with help of Berger titration and/or calculated theoretically. The saturation (SI) index was calculated as:

$$SI = \log(IAP/K) \quad (1)$$

where IAP is the ion activity product and K is the solubility product of calcite ($10^{-8.4}$) (Plummer and Busenberg, 1982).

Cell staining and microscopic analysis

To determine bacterial abundances, 2–5 ml of formaldehyde-fixed samples were stained with DAPI (final concentration, 6.7 mg ml^{-1}) (Porter and Feig, 1980) and filtered onto black polycarbonate filters (pore size 0.22 μm, diameter 25 mm, Osmonics). At least 1000 bacteria per sample were counted by epifluorescence microscopy (AxioImager.M1, filter set 01, Carl Zeiss) at 1000× magnification. For the analysis of bacterial community structure by CARD-FISH, 5–7 ml of formaldehyde-fixed samples were filtered onto white polycarbonate filters (type GTTP, pore size 0.2 μm, diameter 47 mm, Millipore) within 24 h after sampling, air-dried and stored at –20°C for later processing. The embedding of filters in low melting point agarose, the enzymatic predigestion of cell walls and the subsequent staining by CARD-FISH was performed as described previously (Sekar *et al*., 2003). Horseradish peroxidase-labelled oligonucleotide probes were applied targeted to all *Bacteria* (Daims *et al*., 1999), to the *Cytophaga–Flavobacteria* lineage of *Bacteroidetes*, to the α, β and γ subgroups of *Proteobacteria*, to *Actinobacteria* (Amann *et al*., 1995), and to members of the family *Methylophilaceae* (Friedrich *et al*., 2003). In addition, two newly designed probes were used that detect two narrow subgroups within the *Methylophilaceae* (see below for details of probe design and definition of target groups). Signal amplification was performed with tyramides that were custom labelled with the fluorescent dye Alexa488 (Invitrogen). The hybridized filters were counterstained with DAPI (1 mg ml^{-1}) and the relative abundance of hybridized cells in 500–1000 DAPI-stained cells was determined at UV and blue (Zeiss filter set 09) excitation by epifluorescence microscopy.

MAR-FISH

Portions of fixed samples (2.5–5 ml) were filtered onto white polycarbonate membrane filters (pore size 0.2 μm, diameter 25 mm, GTTP, Millipore). Filters were pre-treated and hybridized according to the standard CARD-FISH protocol. Hybridized filter sections were either stored at –20°C or processed immediately for MAR as previously described (Alonso and Pernthaler, 2005; Salcher *et al*., 2008). Manual evaluation of the fractions of MAR-positive cells in all hybridized cells was performed in bright field illumination and blue excitation. At least 500 hybridized cells were inspected per sample.

Cloning and sequence analysis of 16S rRNA genes

DNA was extracted from filters with the PowerSoil DNA Isolation Kit (MOBIO laboratories). For pool NGIII the DNA

recovered from two time points was pooled. Primers GM3F and GM4R (Muyzer *et al.*, 1995) were used for the amplification of 16S rRNA genes. The purified PCR products (QIAquick PCR purification kit, QIAGEN) were cloned into *Escherichia coli* (TOPO TA cloning kit, Invitrogen) and insert-bearing plasmids were prepared with the QIAprep Spin Miniprep Kit (QIAGEN). Sequencing of 16S rRNA genes with M13 vector primers and primer GM1F (Muyzer *et al.*, 1993) was carried out on an ABI 3130x Genetic Analyser using the ABI BigDye chemistry (Applied Biosystems). Partial sequences were assembled (Vector NTI, Invitrogen) and checked for chimeric origin by the software Mallard and Pintail (Ashelford *et al.*, 2005).

Phylogenetic analysis, diversity estimates

The sequences were aligned together with the closest relatives obtained from a BLAST search (Altschul *et al.*, 1990) using the SINA online aligner tool of the SILVA project (Pruesse *et al.*, 2007). The resulting set was merged with the Silva 16S rRNA database (Release 98) using the ARB software package (Ludwig *et al.*, 2004). The alignments of the phylogenetic clusters of interest were subsequently manually optimized. Only sequences > 1200 nucleotides were used for subsequent phylogenetic analysis. Tree reconstruction was first performed with the maximum parsimony (MP) tool of ARB, and subsequent maximum likelihood (ML) analyses were carried out on subsets of 120–600 sequences per group using the RAxML algorithm (Stamatakis *et al.*, 2005) implemented in ARB and group-specific filters (50% variability). Bootstrapped ML trees (100 repetitions) were calculated on a dedicated web server (Stamatakis *et al.*, 2008). These trees were imported into ARB, reduced to a subset of the most closely related sequences, and nodes with bootstrap values < 50% were collapsed into multifurcations.

The diversity of OTUs with similarities ranging from 85% to 99% was estimated for each system using the DOTUR software (Schloss and Handelsman, 2005). The overlap of bacterial OTUs (97% similarity) in the three studied pool was assessed with the software SONS (Schloss and Handelsman, 2006). Sets of trimmed sequences (position 600–1500 in *E. coli*) and standard parameter settings were used for both DOTUR and SONS analyses. The dependence of richness estimates by the Chao 1 parameter on the numbers of sequenced rRNA genes was tested separately for all three systems and for the pooled data set using the online input form developed by Kemp and Aller (2004). The numbers of OTUs with a similarity of 97% as provided by DOTUR were used for this purpose.

Development of FISH probes

The probe design tool of ARB was used to develop probes for two clusters within the family *Methylophilaceae*. Competitors were designed for non-target sequences that featured only a single mismatch at the probe target site. The CARD-FISH procedure with the newly developed probes and competitors was optimized on samples collected from pool NGIII, using hybridization buffers with increasing concentrations of formamide (25% to 70%) to define stringent conditions.

Acknowledgements

We would like to thank the speleologists of the SGH-Bern, especially Martin Schneider and Rolf Siegenthaler for organization of and taking part in sampling procedures, Hans-Peter Kohler (EAWAG) for chemical analysis of water samples, Roland Psenner and Josef Franzoi (University of Innsbruck) for their help in the interpretation of chemical data, and Michaela Salcher and Michael Zeder for discussions and support. This study was supported by a stipend of the Research Commission of the University of Zurich.

References

Alfreider, A., Vogt, C., and Babel, W. (2002) Microbial diversity in an in situ reactor system treating monochlorobenzene contaminated groundwater as revealed by 16S ribosomal DNA analysis. *Syst Appl Microbiol* **25:** 232–240.

Alonso, C., and Pernthaler, J. (2005) Incorporation of glucose under anoxic conditions by bacterioplankton from coastal North Sea surface waters. *Appl Environ Microbiol* **71:** 1709–1716.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Amann, R., and Fuchs, B.M. (2008) Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nat Rev Microbiol* **6:** 339–348.

Amann, R., Ludwig, W., and Schleifer, K. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59:** 143–169.

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J., and Weightman, A.J. (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Appl Environ Microbiol* **71:** 7724–7736.

Bachman, L., Lindsey, B., Brakebill, J., and Powars, D. (1998) Ground-water discharge and base-flow nitrate loads of nontidal streams, and their relation to a hydrogeomorphic classification of the Chesapeake Bay watershed, Middle Atlantic Coast. In US Geological Survey Water-Resources Investigations Report, p. 71.

Balkwill, D.L., Murphy, E.M., Fair, D.M., Ringelberg, D.B., and White, D.C. (1998) Microbial communities in high and low recharge environments: implications for microbial transport in the vadose zone. *Microb Ecol* **35:** 156–171.

Barton, H.A., Taylor, N.M., Lubbers, B.R., and Pemberton, A.C. (2006) DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. *J Microbiol Methods* **66:** 21–31.

Barton, H.A., Taylor, N.M., Kreate, M.P., Springer, A.C., Oehrle, S.A., and Bertog, J.L. (2007) The impact of host rock geochemistry on bacterial community structure in oligotrophic cave environments. *Int J Speleol* **36:** 93–104.

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., *et al.* (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J* **3:** 1093–1104.

Crump, B.C., and Koch, E.W. (2008) Attached bacterial populations shared by four species of aquatic angiosperms. *Appl Environ Microbiol* **74:** 5948–5957.

Daims, H., Brühl, A., Amann, R., Schleifer, K.-H., and Wagner, M. (1999) The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol* **22:** 434–444.

Derakshani, M., Lukow, T., and Liesack, W. (2001) Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Appl Environ Microbiol* **67:** 623–631.

Donachie, S.P., Hou, S., Lee, K.S., Riley, C.W., Pikina, A., Belisle, C., *et al.* (2004) The Hawaiian Archipelago: a microbial diversity hotspot. *Microb Ecol* **48:** 509–520.

Egli, K., Bosshard, F., Werlen, C., Lais, P., Siegrist, H., Zehnder, A.J., and Van der Meer, J.R. (2003) Microbial composition and structure of a rotating biological contactor biofilm treating ammonium-rich wastewater without organic carbon. *Microb Ecol* **45:** 419–432.

Einsiedl, F., Hertkorn, N., Wolf, M., Frommberger, M., Schmitt-Kopplin, P., and Koch, B.P. (2007) Rapid biotic molecular transformation of fulvic acids in a karst aquifer. *Geochim Cosmochim Acta* **71:** 5474–5482.

Fetzner, S., and Lingens, F. (1994) Bacterial dehalogenases: biochemistry, genetics, and biotechnological applications. *Microbiol Mol Biol Rev* **58:** 641–685.

Ford, D.C., and Williams, P. (2007) *Karst Hydrogeology and Geomorphology*. New York, USA: Wiley.

Friedrich, U., Van, L.H., Altendorf, K., and Lipski, A. (2003) Microbial community and physicochemical analysis of an industrial waste gas biofilter and design of 16S rRNA-targeting oligonucleotide probes. *Environ Microbiol* **5:** 183–201.

Griebler, C., and Lueders, T. (2009) Microbial biodiversity in groundwater ecosystems. *Freshwater Biol* **54:** 649–677.

Grimvall, A., and de Leer, E.W.B. (eds) (1995) *Naturally-Produced Organohalogens*. Dordrecht, the Netherland: Kluwer Academic Publisher.

Hama, T., and Yanagi, K. (2001) Production and neutral aldose composition of dissolved carbohydrates excreted by natural marine phytoplankton populations. *Limnol Oceanogr* **46:** 1945–1955.

Häuselmann, P. (2002) Cave genesis and its relationship to surface processes: investigations in the Siebenhengste region (BE, Switzerland). In *Höhlenforschung im Gebiet Sieben Hengste-Hohgant*. Allschwil, Switzerland: Speleo Projects, Nr6.

Hugenholtz, P., Pitulle, C., Hershberger, K.L., and Pace, N.R. (1998) Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol* **180:** 366–376.

Hwang, C., Wu, W., Gentry, T.J., Carley, J., Corbin, G.A., Carroll, S.L., *et al.* (2009) Bacterial community succession during in situ uranium bioremediation: spatial similarities along controlled flow paths. *ISME J* **3:** 47–64.

Kalyuzhnaya, M.G., Bowerman, S., Lara, J.C., Lidstrom, M.E., and Chistoserdova, L. (2006) *Methylotenera mobilis* gen. nov., sp. nov., an obligately methylamine-utilizing bacterium within the family *Methylophilaceae*. *Int J Syst Evol Microbiol* **56:** 2819–2823.

Kalyuzhnaya, M.G., Lapidus, A., Ivanova, N., Copeland, A.C., McHardy, A.C., Szeto, E., *et al.* (2008) High-resolution metagenomics targets specific functional types in complex microbial communities. *Nat Biotechnol* **26:** 1029–1034.

Kemp, P.F., and Aller, J.Y. (2004) Estimating prokaryotic diversity: when are 16S rDNA libraries large enough? *Limnol Oceanogr Methods* **2:** 114–125.

Laiz, L., Groth, I., Gonzalez, I., and Saiz-Jimenez, C. (1999) Microbiological study of the dripping waters in Altamira cave (Santillana del Mar, Spain). *J Microbiol Methods* **36:** 129–138.

Levy, D. (2007) Geochemical trends in selected Lechuguilla Cave pools. *J Cave Karst Stud* **69:** 342–350.

Lindström, E., Forslund, M., Algesten, G., and Bergström, A.-K. (2006) External control of bacterial community structure in lakes. *Limnol Oceanogr* **51:** 339–342.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32:** 1363–1371.

Manda, A.K., and Gross, M.R. (2006) Identifying and characterizing solution conduits in karst aquifers through geospatial (GIS) analysis of porosity from borehole imagery: an example from the Biscayne aquifer, South Florida (USA). *Adv Water Resour* **29:** 383–396.

Matsui, K., Honjo, M., and Kawabata, Z. (2001) Estimation of the fate of dissolved DNA in thermally stratified lake water from the stability of exogenous plasmid DNA. *Aquat Microb Ecol* **25:** 147–171.

Mulec, J. (2008) Microorganisms in hypogeon: examples from Slovenian karst caves. *Acta Carsol* **37:** 153–160.

Musat, N., Werner, U., Knittel, K., Kolb, S., Dodenhof, T., van Beusekom, J.E.E., *et al.* (2006) Microbial community structure of sandy intertidal sediments in the North Sea, Sylt-Romo Basin, Wadden Sea. *Syst Appl Microbiol* **29:** 333–348.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59:** 695–700.

Muyzer, G., Teske, A., Wirsen, C., and Jannasch, H. (1995) Phylogenetic relationship of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gel electrophoresis of 16S rDNA fragments. *Arch Microbiol* **164:** 165–172.

Personne, J.C., Poty, F., Mahler, B.J., and Drogue, C. (2004) Colonization by aerobic bacteria in karst: laboratory and in situ experiments. *Ground Water* **42:** 526–533.

Plummer, L.N., and Busenberg, E. (1982) The solubilities of calcite, aragonite and vaterite in CO_2-H_2O solutions between 0 and 90°C, and an evaluation of the aqueous model for the system $CaCO_3$-CO_2-H_2O. *Geochim Cosmochim Acta* **46:** 1011–1040.

Porter, K., and Feig, Y. (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25:** 943–948.

Pronk, M., Goldscheider, N., and Zopfi, J. (2009) Microbial communities in karst groundwater and their potential use for biomonitoring. *Hydrogeol J* **17:** 37–48.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned

ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188–7196.

Rothberg, J.M., and Leamon, J.H. (2008) The development and impact of 454 sequencing. *Nat Biotechnol* **26:** 1117–1124.

Salcher, M.M., Pernthaler, J., Zeder, M., Psenner, R., and Posch, T. (2008) Spatio-temporal niche separation of planktonic *Betaproteobacteria* in an oligo-mesotrophic lake. *Environ Microbiol* **10:** 2074–2086.

Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71:** 1501–1506.

Schloss, P.D., and Handelsman, J. (2006) Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl Environ Microbiol* **72:** 6773–6779.

Sekar, R., Pernthaler, A., Pernthaler, J., Warnecke, F., Posch, T., and Amann, R. (2003) An improved protocol for quantification of freshwater Actinobacteria by fluorescence in situ hybridization. *Appl Environ Microbiol* **69:** 2928–2935.

Shaw, A.K., Halpern, A.L., Beeson, K., Tran, B., Venter, J.C., and Martiny, J.B. (2008) It's all relative: ranking the diversity of aquatic bacterial communities. *Environ Microbiol* **10:** 2200–2210.

Shigematsu, T., Era, S., Mizuno, Y., Ninomiya, K., Kamegawa, Y., Morimura, S., and Kida, K. (2006) Microbial community of a mesophilic propionate-degrading methanogenic consortium in chemostat cultivation analyzed based on 16S rRNA and acetate kinase genes. *Appl Microbiol Biotechnol* **72:** 401–415.

Simon, K., Pipan, T., and Culver, D. (2007) A conceptual model of the flow and distribution of organic carbon in caves. *J Cave Karst Stud* **69:** 279–284.

Sogin, M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., *et al.* (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103:** 12115–12120.

Stamatakis, A., Ludwig, T., and Meier, H. (2005) RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21:** 456–463.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol* **57:** 758–771.

Tanner, M.A., Goebel, B.M., Dojka, M.A., and Pace, N.R. (1998) Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants. *Appl Environ Microbiol* **64:** 3110–3113.

Tian, Y.J., Yang, H., Wu, X.J., and Li, D.T. (2005) Molecular analysis of microbial community in a groundwater sample polluted by landfill leachate and seawater. *J Zhejiang Univ Sci B* **6:** 165–170.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428:** 37–43.

Wagner, M., and Horn, M. (2006) The *Planctomycetes, Verrucomicrobia, Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr Opin Biotechnol* **17:** 241–249.

Yue, J.C., and Clayton, M.K. (2005) A similarity measure based on species proportions. *Commun Stat Theory Methods* **34:** 2123–2131.

Zhou, J., Gu, Y., Zou, C., and Mo, M. (2007) Phylogenetic diversity of bacteria in an earth-cave in Guizhou province, southwest of China. *J Microbiol* **45:** 105–112.

Manuscript II:

Investigation of bacterioplankton communities in aquatic karst pools in Bärenschacht cave of Bernese Oberland

Extended abstract
Proceedings of 15th International Congress of Speleology, Texas, USA

INVESTIGATION OF BACTERIOPLANKTON COMMUNITIES IN AQUATIC KARST POOLS IN BÄRENSCHACHT CAVE OF BERNESE OBERLAND

TATIANA SHABAROVA, JAKOB PERNTHALER
Limnological Station, Institute of Plant Biology, University of Zurich, Seestrasse 187, CH-8802 Kilchberg, Switzerland

Karst subterraneous aquifers are highly diverse in structure and very important in the formation of ground water, which is the main source of freshwater supply for a significant proportion of the world's population. Microorganisms can play an important role in karstification, carbon cycle and element mobility, but so far little is known about the bacteria of aquatic karst ecosystems. In this study, karst pools with differing hydrology in the Bärenschacht cave of the Bernese Oberland, Switzerland were investigated for a period of six months. Two crystalline pools were supplied by dripping water whereas one epiphreatic pool was renewed only by the rising groundwater table at intervals of several days to months. Chemical parameters such as conductivity, pH, ion concentration, as well as bacterial abundance and diversity were determined at several time points. The investigated pools showed remarkably different physicochemical parameters as well as bacterial properties. Although the dominant bacterial group in all three systems was β-*Proteobacteria*, almost no population overlap inside this group was found between the crystalline pools and the epiphreatic system. *Actinobacteria* were present mainly in the systems with dripping water supply, whereas bacteria from the *Flavobacteriaceae* family were identified in both types of systems. Some microorganisms affiliated with *Bacteroidetes* could also be isolated and investigated in more detail. Generally, most of the identified microorganisms were not closely related to typical freshwater bacteria. Therefore, karst habitats might represent an environment for very specialized microorganisms.

1. Introduction

Ground water represents a major source of freshwater supply for a significant proportion of the world's population, and subterraneous karst aquifers play a vital role in groundwater genesis and initial distribution. Presently, this resource is coming under increasing pressure due to the general recession of glaciers in the context of global warming (Ford and Williams, 2007).

The Bärenschacht cave of the Bernese Oberland (Switzerland) is characterized by an entrance part formed of shafts and steeply dipping galleries of mostly vadose origin which leads into a labyrinth consisting of galleries of mainly phreatic origin. The surface rocks above the cave consist of Globigerina marls and Flysch. Therefore, the water input into Bärenschacht originates almost entirely from the nearby Siebenhengste and Schrattenfluh cave systems that are located in the same catchment area (Häuselmann, 2002). Since Bärenschacht forms the hydrological link between these systems and the springs draining into Lake Thun (Häuselmann, 2002), it allows insight into the genesis and behaviour of an undisturbed deep karst system. Periodic flood events typically result in a rise of the water table in the labyrinthine part of the Bärenschacht cave system. As a consequence, there are numerous relatively small flood-formed pools in this section of the cave with water renewal times ranging from days to years. These epiphreatic pools represent comparatively easily accessible model systems for a better understanding of the chemical and microbiological transformation that may occur in karst ground water at oxic condition.

Microorganisms can play an important role in karstification and they can influence carbon cycle and element mobility (Ford and Williams, 2007, Gabrovsek, *et al*., 2000). So far research about the subsurface microbiota in caves has frequently focused on chemoautotrophy, e.g., their utilization of the available sulphur sources (Macalady, *et al.*, 2008, Barton and Luiszer, 2005). By contrast, little is known about the composition of the heteroorganotrophic microbial assemblages, in particular of those planktonic microbes that inhabit undisturbed and fully oxygenized karst water systems.

We studied the abundances and community composition of planktonic microbes in three oxygenated pools from the epiphreatic and vadose zone of Bärenschacht that differed in hydrology and water chemistry. In one of the systems we moreover compared two molecular biological approaches for cultivation-independent community analysis, 16S rRNA sequences analysis and determinative whole-cell fluorescence *in situ* hybridization (FISH).

2. Materials and Methods

Three pools in the labyrinth part of Bärenschacht cave were sampled at least monthly from October 2007 to March 2008. The epiphreatic pool NGIII is located in the Grand Nord gallery; it is approximately 20 m long and 0.5 m deep. The two pools JI and JII are similar in dimensions as NGIII, but differ in hydrology. They are situated in the Jessica gallery of the upper, non-flooded part of the labyrinth, and both JI and JII are exclusively fed by seepage water. Water samples were collected with sterile syringes from a depth of 5 cm. Subsamples for microscopic analysis were fixed with formaldehyde (2% final concentration) at the site. On several occasions (NGIII: December 29, 2007 and February 20, 2008; JI and JII February 25, 2008) unfixed pool water (1.5 to 2.5 l) was filtered onto white polycarbonate filters (type GTTP: 0.2 μm pore size; diameter, 47 mm, Millipore) in the cave to collect microbial biomass for later DNA extraction. The filters were directly placed into sterile tubes containing lysis buffer of the PowerSoil DNA Isolation Kit (MOBIO laboratories) for transport. All samples were delivered to the lab within 24 h after collection.

A data logger was placed on the bottom of the NGIII pool to monitor water temperature and water level. Conductivity and pH were measured after return to the laboratory with an InoLab benchtop meter (WTW) and a pH-Vision 6071 microcomputer (Jenco Electronics), respectively. Thirty ml of water sample was frozen for late determination of total organic carbon (TOC) on a TOC-5000 analyzer (Shimadzu). Nitrate concentrations in pool water were measured photometrically (cadmium reduction method) (Wood, *et al*., 1967) after filtration through membrane filters (type GTTP, Millipore) that had been pre-rinsed with 2 M HCl and deionised sterile water (Milli-Q, Millipore). The concentrations of sulphate, chloride and of cations were assessed by ion chromatography (Compact IC 761, Metrohm).

Bacterial abundances in pool NGIII were determined from formaldehyde-fixed samples. Portions of 2 to 5 ml were stained with 4'-6-diamidino-2-phenylindole (DAPI, final concentration, 6.7 mg ml^{-1}) (Porter and Feig, 1980) and filtered onto black polycarbonate filters (pore size, 0.22 mm, diameter, 25 mm, Osmonics). At least 1000 bacteria per sample were counted by epifluorescence microscopy (AxioImager.M1, filter set 01, Carl Zeiss) at 1000X magnification. Five to 7 ml of formaldehyde-fixed samples from pool NGIII were filtered onto white polycarbonate filters (type GTTP, 0.2 μm pore size; diameter, 47 mm, Millipore), air-dried and stored at -20° C for later processing. Staining by FISH and catalyzed reporter deposition (CARD) was performed as described previously (Sekar, *et al*., 2003). Horseradish peroxidase labelled oligonucleotide probes were applied that were targeted to all Bacteria (EUB I-III) (Daims, *et al*., 1999), β-*Proteobacteria* (BET42a), members of the Cytophaga-Flavobacteria lineage of Bacteroidetes (CF319a), α-*Proteobacteria* (ALF968), γ-Proteobacteria (GAM42a), and *Actinobacteria* (HGC69a) (Amann, *et al*., 1995). Tyramides custom labelled with the fluorescent dye Alexa488 (Invitrogen) were used for signal amplification. The hybridized filters were counterstained with DAPI (1 mg ml^{-1}) and the fractions of hybridized cells were determined in 500–1000 DAPI-stained cells by epifluorescence microscopy.

Biomass for DNA extraction was collected from JI and JII on Feb 25, 2008. NGIII biomass samples from two time points were pooled (December 29, 2007 and February 20, 2008). DNA was extracted with the PowerSoil DNA Isolation Kit (MOBIO laboratories). Primers GM3F and GM4R (Muyzer, *et al*., 1995) were used for the amplification of 16S rRNA genes. The purified PCR products (QIAquick PCR purification kit, QIAGEN) were cloned into E. coli (TOPO TA cloning kit, Invitrogen) and insert-bearing plasmids were prepared with the QIAprep Spin Miniprep Kit (QUIAGEN). Sequencing of 16S rRNA genes with primers M13R, M13F (Messing, 1983) and GM1 F (Muyzer, *et al*., 1993) was carried out on an ABI 3130x Genetic Analyzer using the ABI BigDye chemistry (Applied Biosystems). Partial sequences were assembled (Vector NTI, Invitrogen) and checked for chimeric origin by the software Pintail (Ashelford, *et al*., 2005). Phylogenetic analyses were performed with the ARB software package (Ludwig, *et al*., 2004).

Cave water samples were plated on three types of media: YST (0.5g starch, 0.5 g yeast extract, 0.5 g peptone), PCA (Merck) and PCA/10. Two dilution factors were used: 10^{-1} ml and 10^{-2} ml. The plates were incubated first for 1 to 3 days at 18° C, and then for several weeks at 8° C. Isolated strains were weekly subcultured and identified by sequencing of their 16S rRNA genes.

3. Results and Discussion

Water temperature and pH were rather similar in all three pools and did not substantially vary during the study period (6.4 to 6.8° C and 7.6 to 7.8, respectively). Four flooding events were recorded, with maximal levels of the water table between 3 m (March 1–3, 2008) and 16 m (October 5–6, 2007)(Fig. 1). The effect of periodic flooding on the water chemistry of the epiphreatic NGIII pool was reflected in higher variability of TOC concentrations and conductivity,

as well as in changes in NO_3 concentrations in time periods separated by flooding events (e.g., 1.2 mg l^{-1} in October/ November 2007 vs. 2.2 mg l^{-1} in December 2007/January 2008). TOC concentration in NGIII (3.2 ± 0.9 mg l^{-1}, mean ± 1 standard deviation) significantly exceeded those of the two crystalline pools (JI: 0.7 ± 0.3 mg l^{-1}; JII: 0.5 ± 0.4 mg l^{-1}), whereas the opposite was true for conductivity (NGIII: 190 ± 20 µS cm^{-1}; JI: 345 ± 5 µS cm^{-1}; JII: 360 ± 2.5 µS cm^{-1}). Concentrations of NO_3 and Ca^{2+} were substantially higher in pool JI (5.46 ± 0.23 mg l^{-1}, and 72 ± 1 mg l^{-1}) than in the other two study systems (JII: 2.89 ± 0.17 mg l^{-1} and 57 ± 1 mg l^{-1}; NGIII: 1.8 ± 0.7 mg l^{-1} and 41 ± 4 mg l^{-1}). By contrast, JII featured substantially higher concentrations of SO_4^{2-} and Mg^{2+} than the other two pools.

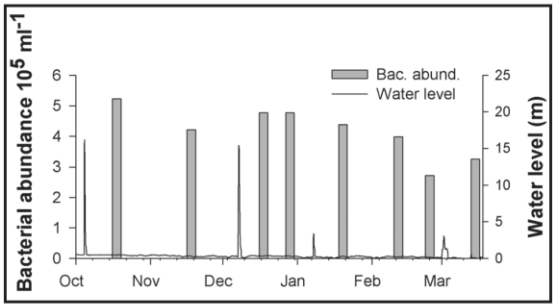


Figure 1: Changes in water level and of total bacterial abundance observed at the NGIII pool of Bärenschacht cave between October 2007 and March 2008.

Bacterial abundances in 8 samples from pool NGIII ranged between 2.7×10^5 ml^{-1} and 5.2×10^5 ml^{-1} (Fig. 1). Total cell numbers tended to declined with time if the period between successive flooding events exceeded one month (October to November 2007 and January to March 2008). This indicates that besides a growth arrest due to substrate limitation mortality by protistan predation or viral lysis might have affected the microbial communities (Pernthaler, 2005). However, a microscopic inspection of several samples suggested that the abundances of free-living phagotrophic flagellates were very low (data not shown).

No quantitative microscopic analysis of bacteria on membrane filters was possible in samples from JI and JII, because (i) microbial cell numbers in the other two pools were extremely low (<10^4 cells ml^{-1}) and (ii) the water from these systems contained high numbers of inorganic particles that impeded the evaluation of more concentrated samples. Therefore, no precise determination of bacterial abundances could be achieved. In order to quantify microbes from crystalline subsurface water pools it might be necessary to separate cells from particles by density gradient centrifugation (Fazi, *et al.*, 2005), or to perform analyses by flow cytometry (Hammes, *et al.*, 2008). However, the low cell numbers in JI and JII might pose a problem even for these approaches.

Created libraries of bacterial 16S rRNA genes consisted of 102 clones for NGIII, 108 for JI, and 103 for JII. A large fraction of sequence types from NGIII and JII (43 and 45%, respectively) was related to β-*Proteobacteria*, but <15% of all sequences from JI. A more detailed analysis of the phylogenetic affiliation of β-*Proteobacteria* revealed striking differences between the three systems (Fig. 2): While *Methylophilaceae* represented the dominant group in NGIII, these bacteria were entirely absent in the other two habitats. Bacteria from this linea have been found in high abundances in the sub-to anoxic layers of a mesotrophic lake, whereas they were rare in oxygenated wat (Salcher, et al., 2008). Most β-*Proteobacteria* in the pools from the vadose zone were affiliated with either *Comomonadaceae* (JI) or *Oxalobacteriaceae* (JII). β-*Proteobacteria* are known to be a major component of bacterioplankton in many surface freshwater systems (Glöckner, *et al.*, 1999). However, none of the sequence types from the studied pools was closely related to the typical lineages that are known from lakes or rivers, suggesting that subsurface karst pools may harbour a highly specialized planktonic microflora. This conclusion is further corroborated by the phylogenetic affiliation of other sequence types in our libraries (data not shown). For example, several sequences from NGIII and JII fell into the TM7 and OP10 phyla, respectively, that are not known to occu in surface freshwater habitats (Zwart, *et al.*, 2002). The crystallir pool with highest NO_3 concentration (JI) harboured several phylotypes affiliated with *Nitrospira*, a group of nitrite oxidizing bacteria known from soils (Roesch, *et al.*, 2007) and nitrifying biofilms (Daims, *et al.*, 2001). These bacteria were

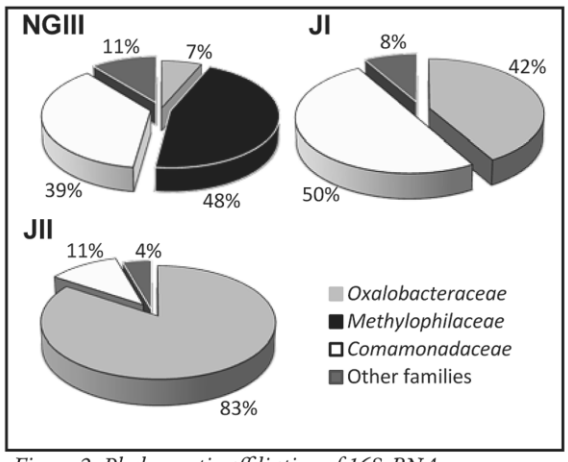


Figure 2: Phylogenetic affiliation of 16SrRNA gene sequences of β-Proteobacteria in the three studied pools NGIII, JI, and JII.

also detected in karst spring water samples during periods of low input of agriculturally influenced surface water (Pronk, *et al*., 2009). Interestingly, some phylotypes affiliated with *Bacteroidetes* were present both in the clone libraries and in the culture collection, which is in contrast to findings from surface waters (Alonso, *et al*., 2007).

The detection rates of bacteria in NGIII by CARD-FISH varied from <40% of total cell counts in October/November 2007 to >60% in January/February 2008 (mean, 54 ± 12%). This lies within the range of values reported for more productive surface freshwater systems (Posch, *et al*., 2009). Only between 49 % and 75 % of all hybridized *Bacteria* could be further identified by the set of group-specific probes, confirming the presence of other bacterial lineages as indicated by sequence analysis. β-*Proteobacteria* represented the most abundant bacterial group that could be identified by FISH in all samples from NGIII (48 ± 7% of *Bacteria*), followed by microbes affiliated with the *Cytophaga-Flavobacteria* lineage of *Bacteroidetes* (CFB, 8 ± 2% of *Bacteria*). By contrast the relative abundances of α-*Proteobacteria* never exceeded 6% of hybridized cells, and *Actinobacteria* could only be detected in 2 out of 8 samples. Figure 3 compares the fractions of sequence types affiliated with different bacterial taxa in the clone library from NGIII with the proportions of these groups detected by FISH on the same sampling time points. This allows an assessment of the bias that is potentially introduced by a preferential PCR amplification of particular sequence types (von Wintzingerode, *et al*., 1997). A clear overrepresentation of phylotypes affiliated with CFB was observed. By contrast, a somewhat higher fraction of β-*Proteobacteria* was found in pool water by direct microscopic inspection than by cloning of 16S rRNA genes. Besides differential amplification by PCR it is conceivable that CFB might have featured higher numbers of rRNA operons than other bacteria. Bacteria affiliated with CFB formed the most prominent cultivable bacterial group from NGIII, and it has been shown that isolates with higher rRNA operon number tend to form colonies on solid media more rapidly than others (Klappenbach, *et al*., 2000).

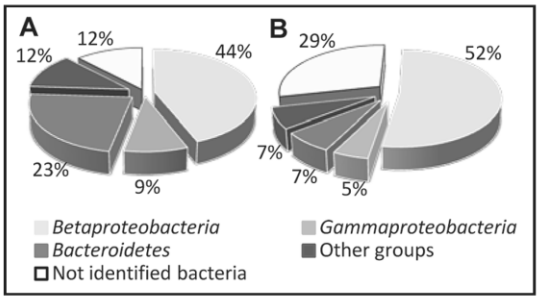


Figure 3:. Community composition of Bacteria in pool NGIII according to phylogenetic analysis of 16S rRNA gene sequences (A) or microscopic assessment after CARD-FISH staining (B).

4. Conclusions

Three fully oxygenated subsurface karst pools with contrasting hydrogeology and chemistry were found to harbor diverse microbial assemblages. β-*Proteobacteria* constituted the most prominent bacterial lineage in all three systems, but each habitat featured a set of distinct 16S rRNA phylotypes affiliated with this group. Moreover, the fine-scale phylogenetic composition of the microbial assemblages in the studied pools clearly differed from surface water communities. Therefore, subsurface karst pools appear to feature a highly specialized planktonic microflora. However, a direct microscopic analysis of bacterial community composition in subsurface karst pools with long water residence times may be hampered by both high background of particulate matter and low cell numbers

5. References:

ALONSO, C., F. WARNECKE, R. AMANN, and J. PERNTHALER (2007) High local and global diversity of Flavobacteria in marine plankton. Environmental Microbiology 9, 1253–1266.

AMANN, R.I., W. LUDWIG, and K.H. SCHLEIFER (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiological Reviews 59, 143–169.

ASHELFORD, K.E., N.A. CHUZHANOVA, J.C. FRY, A.J. JONES, and A.J. WEIGHTMAN (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Applied and Environmental Microbiology 71, 7724–7736.

BARTON, H.A. and F. LUISZER (2005) Microbial metabolic structure in a sulfidic cave hot spring: Potential mechanisms of biospeleogenesis. Journal of Cave and Karst Studies 67, 28–38.

DAIMS, H., A. BRUHL, R. AMANN, K.H. SCHLEIFER, and M. WAGNER (1999) The domain–specific probe EUB338 is insufficient for the detection of all Bacteria: Development and evaluation of a more comprehensive probe set. Systematic and Applied Microbiology 22, 434–444.

DAIMS, H., J.L. NIELSEN, P.H. NIELSEN, K.H. SCHLEIFER, and M. WAGNER (2001) In situ characterization of Nitrospira–like nitrite oxidizing bacteria active in wastewater treatment plants. Applied and Environmental Microbiology 67, 5273–5284.

FAZI, S., S. AMALFITANO, J. PERNTHALER, and A. PUDDU (2005) Bacterial communities associated with benthic organic matter in headwater stream microhabitats. Environmental Microbiology 7, 1633–1640.

Ford, D.C. and P. Willams (2007) Karst Hydrogeology and Geomorphology. John Wiley & Sons, London, 576 p.2nd ed.

GABROVSEK, F., B. MENNE, and W. DREYBRODT (2000) A model of early evolution of karst conduits affected by subterranean CO2 sources. Environmental Geology 39, 531–543.

GLÖCKNER, F.O., B.M. FUCHS, and R. AMANN (1999) Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. Applied and Environmental Microbiology 65, 3721–3726.

HAMMES, F., M. BERNEY, Y.Y. WANG, M. VITAL, O. KOSTER, and T. EGLI (2008) Flow–cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. Water Research 42, 269–277.

Häuselmann, P. (2002) Cave genesis and its relationship to surface processes: Investigations in the Siebenhengste region (BE, Switzerland). PhD thesis, University of Fribourg, Fribourg.

KLAPPENBACH, J.A., J.M. DUNBAR, and T.M. SCHMIDT (2000) rRNA operon copy number reflects ecological strategies of bacteria. Applied and Environmental Microbiology 66, 1328–1333.

LUDWIG, W., O. STRUNK, R. WESTRAM, L. RICHTER, H. MEIER, YADHUKUMAR, A. BUCHNER, T. LAI, S. STEPPI, et al. (2004) ARB: a software environment for sequence data. Nucleic Acids Research 32, 1363–1371.

MACALADY, J.L., S. DATTAGUPTA, I. SCHAPERDOTH, D.S. JONES, G.K. DRUSCHEL, and D. EASTMAN (2008) Niche differentiation among sulfur–oxidizing bacterial populations in cave waters. Isme Journal 2, 590–601.

MESSING, J. (1983) New M13 Vectors for Cloning. Methods in Enzymology 101, 20–78.

MUYZER, G., A. TESKE, C.O. WIRSEN, and H.W. JANNASCH (1995) Phylogenetic relationship of Thiomicrospira species and their identification in deep–sea hydrothermal ventes by denaturing gradient gel electrophoresis of 16S rDNA fragments. Archives of Microbiology 164, 165–172.

MUYZER, G., E.C. DE WAAL, and A.G. UITTERLINDEN (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction–amplified genes coding for 16S rRNA. Applied and Environmental Microbiology 59, 695–700.

PERNTHALER, J. (2005) Predation on prokaryotes in the water column and its ecological implications. Nature Reviews Microbiology 3, 537–546.

PORTER, K.G. and Y.S. FEIG (1980) The use of DAPI for identifying and counting aquatic microflora. Limnology and Oceanography 25, 943–948.

POSCH, T., J. FRANZOI, M. PRADER, and M. M. SALCHER (2009) New image analysis tool to study biomass and morphotypes of three major bacterioplankton groups in an alpine lake. Aquatic Microbial Ecology 54, in press.

PRONK, M., N. GOLDSCHEIDER, and J. ZOPFI (2009) Microbial communities in karst groundwater and their potential use for biomonitoring. Hydrogeology Journal 17, 37–48.

ROESCH, L.F., R.R. FULTHORPE, A. RIVA, G. CASELLA, A.K.M. HADWIN, A.D. KENT, S.H. DAROUB, F.A.O. CAMARGO, W.G. FARMERIE, et al. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. Isme Journal 1, 283–290.

SALCHER, M.M., J. PERNTHALER, M. ZEDER, R. PSENNER, and T. POSCH (2008) Spatio–temporal niche separation of planktonic Betaproteobacteria in an oligo–mesotrophic lake. Environmental Microbiology 10, 2074–2086.

SEKAR, R., A. PERNTHALER, J. PERNTHALER, F. WARNECKE, T. POSCH, and R. AMANN (2003) An improved protocol for the quantification of freshwater actinobacteria by fluorescence in situ hybridization. Applied and Environmental Microbiology 69, 2928–2935.

VON WINTZINGERODE, F., U.B. GÖBEL, and E. STACKEBRANDT (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR–based rRNA analysis. FEMS Microbiology Reviews 21, 213–229.

WOOD, E.D., F.A.J. ARMSTRONG, and F.A. RICHARDS (1967) Determination of Nitrate in Sea Water by Cadmium–Copper Reduction to Nitrite. Journal of the Marine Biological Association of the United Kingdom 47, 23.

ZWART, G., B.C. CRUMP, M. AGTERVELD, F. HAGEN, and S.K. HAN (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquatic Microbial Ecology 28, 141–155.

3. Transformation of microbial and chemical parameters during stagnation periods in epiphreatic pools

Manuscript III

Mass effects meet species sorting: transformations of microbial assemblages in epiphreatic subsurface karst water pools.

Environmental Microbiology, In revision

Mass effects meet species sorting: transformations of microbial assemblages in epiphreatic subsurface karst water pools.

Tanja Shabarova[1], Franco Widmer[2], & Jakob Pernthaler[1]*

[1]*Limnological Station, Institute of Plant Biology, University of Zurich, Seestr. 187, CH-8802 Kilchberg, Switzerland*

[2]*Molecular Ecology, Reckenholz-Tännikon Research Station ART, Reckenholzstrasse 191, CH-8046 Zurich, Switzerland*

Running head: microbes in subsurface karst water pools

*corresponding author: Phone: 0041 634 92 10; Fax: 0041 634 92 25; Email: pernthaler@limnol.uzh.ch

Abstract:

We investigated the transformations of the microbial communities in epiphreatic karst water pools with different flooding frequencies in a Swiss cave system during the winter season. Molecular fingerprinting of 16S rRNA genes by terminal restriction fragment length polymorphism (T-RFLP) was combined with microscopic and sequence analysis in order to assess if the source water would transport comparable microbial inocula into the pools at consecutive flood events, and if there was an effect of residence time on the microbial assemblages in the pools during the stagnant periods between floodings. Variability in the concentrations of dissolved organic carbon and of conductivity indicated differences between floods as well as changes of pool water over time. β-*Proteobacteria* affiliated with *Methylophilaceae* and *Comamonadaceae* were introduced into the pools in high densities during subsequent flood events. The former lineage was able to persist for extended time periods in the pools; the latter one exhibited considerable microdiversification, as reflected by operational taxonomic units (99 % similarity) that consisted of several discernible genotypes. These β-*Proteobacteria* might, therefore, represent core microbial groups in karst water at least during winter. Within the remaining bacterial taxa, a distinct decrease in estimated total diversity was apparent after only a few weeks of residence; some genotypes were specifically favored by stagnant conditions, whereas a larger proportion was rapidly outcompeted. Thus, the microbial communities could be divided into different components that appeared to be governed by complementary mechanisms of community assembly (dispersal vs. environmental filtering) upon introduction into the pools. High overlap of both, temporary and persistent community members between samplings from two winters, moreover, reflected the seasonal recurrence of the microbial assemblages in the studied karst water habitat.

Introduction

Karst is a landscape that develops on soluble rocks such as carbonate and gypsum. It covers approximately 10 % of the Earth's mainland. Due to the dissolution processes that play a central role in karst genesis, the formation of a subsurface network of conduit aquifers is typical for karst areas and explains the ability of this landscape type to conduct and store precipitation water. Karst water management has received increasing attention during the past decades as a consequence of the increasing pressure on global drinking water resources (Ford and Williams, 2007).

Certain aspects of the microbiology of karst water have increasingly come into focus, specifically the transport and control of pathogens (Pronk et al., 2007; Reischer et al., 2008; Butscher et al., 2011), or microbial degradation of organic contaminants (Iker et al., 2010). However, besides the occurrence, identity and function of particular autotrophs (Barton and Luiszer, 2005; Macalady et al., 2008), relatively little is known about microbes in pristine karst water systems. The existence of autochthonous microbial endokarst communities has been proposed (Farnleitner et al., 2005) that may, however, mainly consist of genotypes also known from other subsurface or surface habitats (Griebler and Lueders, 2009; Shabarova and Pernthaler, 2010). The lion's share of data about karst water microbiota originates from studies of springs and boreholes, which, e.g., allow assessing the effects of seasonality, discharge dynamics, and geographic differences on microbial community composition (Farnleitner et al., 2005; Pronk et al., 2006; Wilhartitz et al., 2009). Such studies typically do not consider the potential community changes that may occur during residence within karst ecosystems or during their transfer through the subsurface realms.

Caves are natural karst formations that may serve as possible access points to karst aquifers between the sink and the spring, thereby providing an additional platform for research on karst water microbiology. Caves specifically provide the opportunity to explore the epiphreatic zone of the karst environment, which can be either defined as the zone between the highest and lowest piezometric water table in the karst massive (Gams, 2003) or as the zone that is flooded regularly (Häuselmann, 2002). Besides its function as conduit for flood water, the epiphreatic environment often exhibits some water storage capacity, as the unevenness of the bedrock floor may lead to the formation of pools of various sizes and shapes. This latter property is, e.g., reflected in an initial rise of conductivity in spring water during flood events that is indicative of 'old' stored water (Ford and Williams, 2007; Ravbar et al., 2011) and that in turn points to a change of water chemistry during its residence in the endokarst environment. By contrast, the concentrations of dissolved organic carbon (DOC) typically decrease during the passage of water through subsurface karst formations (Simon et al., 2007; Simon et al., 2010), which -besides other possible causes, such as a sorption to

mineral surfaces- might also be related to the activity of heteroorganotrophic microorganisms.

Epiphreatic water pools that are accessible to sampling may serve as model habitats to study the potential transformations of the planktonic microbial assemblages during their endokarst residence. Although these pools provide highly stable conditions for bacterial growth for a limited time period, they are intrinsically temporary habitats in that they experience complete water exchange during flooding events. We have previously shown that the planktonic microbial communities in such pools may be surprisingly diverse and that a bacterial population in a stagnant pool could persist almost unchanged over periods of weeks (Shabarova and Pernthaler, 2010). Due to the vertical location of individual pools within the epiphreatic zone and their hydrological connection to the karst water source they typically experience different flooding frequencies. This allows for a more detailed investigation of the respective effects of water residence time, source water inoculum and local (within-pool) variability on the composition and temporal development of the microbial assemblages.

We investigated changes of water chemistry, bacterial abundances, diversity and community composition in three epiphreatic pools that are situated in close vicinity within a single karst cave system and that are flooded at different intervals by the same water source (Fig. 1). We specifically tested the hypothesis that (i) the microbial assemblages in the source water would be similar during repeated flood events within a single season and (ii) that the microbial assemblages in the pools would transform in a parallel manner during the stagnant periods between flooding events, i.e. community changes were mainly determined by water residence time rather than by differences between individual pools and floods.

Results

Water level and temperature monitoring

During the study period water temperature ranged between 7.4 and 7.7 °C in all three pools. The last flood event before the sampling campaign (5.7 m above the regular level) was registered in pool LC on November 18, 2009; no more fluctuation in water level was observed in this system until the end of March (Fig. 1). Pool NGIII was flooded on December 11 2009 (2.2 m) and on December 30 (0.4 m). The water table of the Nord Sump was elevated during some months before the study. It sank to its regular level on December 20, 2009, and rose again between December 25, 2009 and January 9, 2010 (28.5 m). The last flood during the study took place in pool NS on March 2, 2010 (0.4 m).

Chemical parameters

A rise in the concentrations of most ions as well as of conductivity was observed in all pools after the flood events. In pool LC, the concentration of Ca ions (the most abundant cation in all three pools) and conductivity increased between the beginning and the end of the study, by 1.4 times (from 38 to 58 mg l^{-1}) and by 1.3 (from 183 to 240 µS cm^{-1}), respectively (Fig. 2). In pool NGIII the increase was 1.5 times for Ca (from 34.9 to 53.8 mg l^{-1}) and 1.25 times for conductivity (from 172 to 215 µS cm^{-1}). The Ca concentration and conductivity in the final sample of pool NS were 51.9 mg l^{-1} and 222 µS cm^{-1}, and the changes in Ca concentration in this hydrologically unstable pool were significantly correlate with conductivity (R^2=0.7, $P < 0.05$). Concentrations of most of the other ions were under the detection limit, particularly in the earlier NGIII and LC samples. An amount of 10 mg l^{-1} of sulfate was found in the final LC sample, of 7 mg l^{-1} in the final NGIII sample and of 8 mg l^{-1} in the final NS sample. In earlier samples from pools NGIII and NS and in the first sample from pool LC sulfate concentrations were lower than 5 mg l^{-1}, whereas 6 mg l^{-1} were detected in LC samples 2-5. Silica concentrations followed the same trend, i.e., an accumulation over time, in LC from 2.5 to 3.2 mg l^{-1} (1.3 times) and in NGIII from 2.8 to 3.5 mg l^{-1} (1.3 times). The differences between total organic carbon (TOC) concentrations in unfixed and HCl fixed samples were ≤0.2 mg C l^{-1}. Only small differences between DOC and TOC concentrations (≤ 0.3 mg l^{-1}) were observed in all samples. The development of organic carbon concentration over time showed an opposite trend to most other chemical parameters, e.g. DOC decreased from 2.4 to 2.0 mg C l^{-1}in LC and from 2.8 to 1.6 mg C l^{-1} in NGIII (Fig. 2). Values of total and dissolved N (TN, DN) and nitrate were very similar to each other. No changes of nitrate concentration, TN and DN was detected in NGIII (0.42 ± 0.04 mg N l^{-1}, mean ± 1 standard deviation), whereas these values increased from 0.4 to 1 mg N l^{-1} in pool LC. Concentrations of dissolved and total P were below detection limits in all but the final samples from pools NGIII (8.3 and 17.4 µg P l^{-1}) and LC (7.2 and 14.3 µg P l^{-1}).

Abundances of total bacteria and of specific groups

Bacterial abundances in the epiphreatic pools ranged between 1.0×10^5 ($\pm1.0\times10^2$) cells ml^{-1} (sample NS 1) and 2.9×10^5 ($\pm0.2\times10^4$) cells ml^{-1} (sample LC 4) (Fig. S1). In pool NGIII an initial increase of microbial numbers was observed between sample 1 and sample 3 (from 1.9×10^5 $\pm0.2\times10^5$ cells ml^{-1} to 2.6×10^5 $\pm1.7\times10^4$ cells ml^{-1}), followed by a steep drop in sample 4 after the flood on 30.12.2009 (1.8×10^5 $\pm7\times10^3$ cells ml^{-1}). A second slight rise in bacterial numbers was detected in sample 5 (2.1×10^5 $\pm5.4\times10^3$ ml^{-1}). Subsequently, microbial abundances gradually decreased, to 1.2×10^5 ($\pm1.6\times10^4$) cells ml^{-1} in the last sample.

Bacterial numbers also increased in the first 4 samples of pool LC, to a maximum of 2.9×10^5 ($\pm5\times10^4$) cells ml^{-1}. A decline of bacterial numbers was again observed in the later samples from this pool (sample 7: 1.1×10^5 $\pm1.5\times10^4$ cells ml^{-1}). Microbial abundances in pool NS did not exceed 1.4×10^5 cells ml^{-1} during the flood events but reached at least 2×10^5 cells ml^{-1} immediately thereafter (samples 5, 6 and 7).

Actinobacteria and members of the α and γ subgroups of *Proteobacteria* were only present in low proportions (<2.0 %) in all samples; thus no quantification of the abundances of these groups was attempted. β-*Proteobacteria* were the most abundant bacteria identified by fluorescence in situ hybridization (FISH) at all time points in all systems (from 16 % of total cells in pool NS, sample 1 to 46 % in pool LC on the same date); the variability of replicate determinations with probe BET42a was small (±3 %). The temporal development of numbers followed the trend of total bacterial abundances (Figs. 3, S1). Members of the β-*proteobacterial* family *Methylophilaceae* were detected in all samples in proportions ranging from 2 % to 8 % (pool NS, samples 5 and 7, respectively, Fig. 3). Members of the *Cytophaga-Flavobacteria* lineage of *Bacteroidetes* (CFB) as detected by probe CF 319a was the second most abundant large phylogenetic group (from 3 % in sample 1 of pool NS to 10 % in pool NGIII, sample 6).

Analysis of terminal restriction fragment length polymorphisms (T-RFLP)

13 samples and a triplicate control sample from a different season were analyzed by T-RFLP using two restriction enzymes. Panels constructed using all 16 electropherograms revealed a total of 44 and 44 peaks for *Rsa*I and *Msp*I digestions, respectively. Signals presumably originating from β-*Proteobacteria* (as predicted from the *in silico* distribution of T-RFs from a previously established 16S rRNA gene clone library from pool NGIII) (Shabarova and Pernthaler, 2010) were dominating all samples. They represent from 44 % to 88 % of total peak height in the chromatograms of samples after *Msp*I digestions (samples NS 7; NGIII 2) and from 20 to 77 % in digestions with *Rsa*I (samples NS 7; NGIII 3). The difference in the proportions of predicted β-*proteobacterial* T-RFs (i.e., their contribution to total peak height) in a triplicate determination were 15 % and 11 % for *Msp*I and *Rsa*I, respectively. Analysis of community similarities (Pearson method) of the complete T-RFLP data sets did not reveal clear patterns related to either residence time, pool or source water (data not shown). Thus, in order to specifically assess changes within the remaining signals of the chromatograms, all peaks assigned to β-*Proteobacteria* were excluded, as identified by the respective lengths of β-*proteobacterial* T-RFs produced from the clone libraries ± 7 bp (Hartmann and Widmer, 2008) (*Commanadaceae:* 5 peaks at 486-497 bp for *Msp*I and 8

166 peaks at 424-428 bp for *Rsa*I; *Methylophilaceae* 1 peak at 468-480 bp for *Rsa*I). Cluster
167 analysis of this reduced data set yielded comparable dendograms for both enzymes, as
168 established by a Mantel test (two sided P-value=0.001). Based on this analysis it was
169 possible to generally distinguish between two groups of communities in dependence of the
170 water residence time (early vs. late post-flooding assemblages, Fig. 4). The only exception to
171 this pattern (in analyses with both enzymes) was sample NGIII 5 that clustered to the early
172 samples despite a residence time of 14 days. Moreover, the last sample from pool NS
173 clustered separately from all others using enzyme *Rsa*I.

174

175 *16S rRNA gene clone libraries*

176 After the removal of putative chimera, 301 and 304 partial 16S rRNA gene sequences
177 were obtained from samples NGIII 1 and NGIII 3, respectively. As expected, the two
178 microbial assemblages were overlapping to a considerable extend, which was reflected by
179 19 shared operational taxonomic units (OTUs), and in particular by a θ value of 0.8.
180 However, an analysis of predicted diversity vs. sample size (i.e., the stability of Chao 1
181 values in differently sized subsamples) also revealed a remarkable reduction of estimated
182 community diversity with prolonged water residence time (Fig. 5A). 85 of 107 detected OTUs
183 (at 99 % similarity level) in NGIII 1 were represented by only one sequence. The total amount
184 of OTUs in NGIII 3 decreased to 85, of which 56 were singletons. Relative abundance of
185 sequences related to β-*Proteobacteria* ranged from 56 % in NGIII 1 to 70 % in NGIII 3. No
186 pronounced decline in diversity (Fig. 5B) or differences in OTU composition could be
187 observed within this group. β-*Proteobacteria* were mainly composed of members of the
188 *Comamonadaceae* (total: 6 OTUs; NGIII 1: 4 OTUs and 32 % of sequences; NGIII 3: 5 OTUs
189 and 32 % of sequences; Fig. 5C) and of *Methylophylaceae* (total: 5 OTUs; NGIII 1: 4 OTUs
190 and 20 % of sequences; NGIII 3: 4 OTUs and 30 % of sequences). By contrast, other
191 putatively methylotrophic bacteria of the family *Methylococcoceae* (γ-*Proteobacteria*) were
192 found in the first post-flooding sample only (7 % of sequences); these sequences grouped
193 into 3 OTUs that were phylogenetically closely related to the freshwater aerobic methane
194 oxidizer *Crenothrix polyspora* (Fig. 6). The relative abundance of sequence types related to
195 CFB ranged from 15 % in NGIII 1 to 11 % in NGIII 3. Some OTUs from this lineage moreover
196 showed clear successional pattern (Fig. S2), e.g., OTUs affiliated with the genus
197 *Flavobacterium* were dominating in NGIII 1 and were replaced in NGIII 3 by members of the
198 families *Cryomorphaceae* (*Fluviicola* group) and *Flexibacteriaceae.*

199 In order to assess which OTUs were present in the karst system over successive
200 years, additional analyses were performed by including a clone library established from pool

NGIII in winter 2007/2008 (Shabarova and Pernthaler, 2010). More diversification was detected in this pooled data set for several bacterial groups, e.g. there were 9 rather than 6 OTUs of *Comamonadaceae*. 23 OTUs containing more than two thirds of all sequences (526 of 769) were shared between both winters. Only 14 of these interannually occurring OTUs were also shared between samples NGIII 1 and NGIII 3. OTUs common in both winters were related to the *Comamonadaceae* (5 OTUs), *Methylophylaceae* (5 OTUs) and 2 other groups of β-*Proteobacteria* (2 OTUs), to *Methylococcaceae* (2 OTUs) and *Oceanospirillum* (2 OTUs) (γ-*Proteobacteria*), and to the generea *Flavobacterium* (4 OTUs), *Arcicella* (1 OTU), *Fluvicola* (1 OTU) and *Sediminibacter* (1 OTU) of *Bacteroidetes*.

Discussion:

Dominant community members in karst water pools

β-*Proteobacteria* were the dominating group in all ponds at all sampling time points and according to all three molecular approaches (FISH, sequencing, T-RFLP). Similar temporal trends for β-*Proteobacteria* were observed with both FISH and T-RFLP in most samples of pools LC and NGIII, e.g., with respect to the time points of maxima and minima. There were, however, considerable discrepancies between the proportions of β-*Proteobacteria* as estimated either by FISH or by PCR based methods, the latter being systematically higher by 15 to 38 % even in samples with a comparable trend. This is probably a result of disproportional PCR amplification of sequences related to *Methylophilaceae*; the relative abundance of these bacteria as derived from T-RFLP / cloning were 2-10 times (corresponding to 10 % - 31 %) higher than determined by FISH. The largest differences -up to 48 %- were noted in samples taken during the flood events or at the very end of stagnant periods (in particular samples NS 1–3, LC 7 and NGIII 7), which might possibly be due to an underestimation of small and starving cells by FISH (Hoshino et al., 2008).

Sequences or T-RFs related to *Comamonadaceae* represented the most common β-*Proteobacteria* in all samples, except for sample NS 7, where the summed intensities of T-RFs of predicted members of the *Methylophilaceae* were higher. This agrees with the highest abundance of *Methylophilaceae* in this sample detected by FISH. In both bacterial groups, one large OTU was accompanied by several small ones. A conspicuous feature of the *Comamonadaceae* OTUs was the high degree of microdiversification, as indicated by different closest neighbors (Fig. 5C). This suggests considerable niche diversity for these bacteria in the endokarst environment, either due to their metabolic variability (Jaspers et al., 2001; Hahn and Pockl, 2005) or spatial separation (Jezbera et al., 2011). Most of these

bacterial taxa should not be regarded as exclusively 'autochthonous' endokarst microorganisms, since many OTUs were most closely associated to genotypes from surface aquatic habitat rather than karst water: Only two of the 4 closest relatives of the most abundant OTU, AM991232 and AM991264 (accessions B and D in Fig. 5C) originate from karst spring water in the Swiss Alps (Pronk et al., 2009).The second most frequent OTU of *Comamonadaceae* (closest matches: accessions E, F in Fig. 5C) was affiliated with bacteria from surface freshwaters (Salcher et al., 2008; Hahn et al., 2010); its closest cultured relative (98 % similarity), *Limnohabitans curvus,* was isolated from lake water.

Community changes related to residence time and flooding events

For the analysis of the potential effects of flooding event and residence time (Fig. 4), all peaks were removed from the T-RFLP chromatograms that were considered to originate from β-*Proteobacteria*. This was done after observing that the individual determinations from a replicate sample (NGIII 0) did not cluster together without such pre-filtering. It is conceivable that the relatively high size variability of the dominant, presumably β-proteobacterial T-RFs suppressed the significance of the other T-RFs. The diversity in the ponds was undersampled by sequencing (Shabarova and Pernthaler, 2010) (Fig. 5A); thus our *in silico* analysis might have overlooked some peaks originating from β-*proteobacteria*, or genotypes affiliated with other bacterial classes might have produced T-RFs at positions considered as diagnostic for β-*Proteobacteria*. However, since the estimated proportions of β-*Proteobacteria* were comparable in the clone libraries and T-RFLP chromatograms (as deduced from their combined relative peak height), it is likely that our peak filtering approach, by and large, mainly removed bacteria from this subphylum.

T-RFLP analysis of the remaining microbial assemblages with both restriction enzymes yielded a general distinction of the remaining microbial assemblages during and shortly after flooding vs. after prolonged residence that was largely independent of the sampled pool or the individual flooding events (Fig. 4). This might be interpreted as an indication for strong environmental filtering of many non-dominant community members (Stegen et al., 2012), i.e. only a fraction of the newly introduced genotypes could successfully establish themselves in the pools. A similar conclusion is reached from a comparison of the two clone libraries (Fig. 5A): Although both libraries are clearly undersampled (Fig. 5A), they nevertheless reflect a substantial loss of the total estimated microbial diversity in pool NGIII during the stagnant period. Additionally, a more detailed analysis of *Bacteroidetes* related to CFB yielded several examples for shared OTUs that predominantly consisted of sequences from one of the two libraries (Fig. S2).

We also noted differences between flood events in their effects on community development. Specifically, the second flood in pool NGIII did not appear to immediately induce changes of the microbial assemblages, and the total bacterial abundances between samples 4 (1 week after flood) and 5 (2 weeks after flood) only marginally increased (Fig. S1). This might be related to the typical dynamics of flooding events in karst systems: In the initial phase, mainly 'old' water stored within the karst system is transported (Ravbar et al., 2011), as indicated by lower DOC values and high conductivity, whereas the later period is characterized by water with increased DOC concentrations and lower conductivity (i.e., of surface origin). These changes are clearly visible in the water chemistry of pool NS (samples 3-5, Fig. 2). Elevated conductivity values in sample NGIII 4 suggest that the second flooding event also largely consisted of 'old' water. However, if such 'old' water originated from pools comparable to the studied ones, it appears contradictory that the microbial assemblages in the subsequent sample 5 from pool NGIII clustered with others that were obtained during or shortly after flooding with water of presumably surface origin (Fig. 4). Possible explanations might be: (i) the transport process itself represented a type of environmental filtering, e.g. by selective mortality or retention on surfaces, or (ii) bacteria (that are not β-*Proteobacteria*) in flood water mainly originated from habitats other than cave pools.

The last flooding event in pool NS was observed approximately 2 weeks prior to sampling and lead to an incomplete water exchange (only 40 cm above normal water levels; average pool depth: 1.5 m; Fig. 1). Although this sample did cluster with the other 'late' samples in T-RFLP analysis with one enzyme (*Msp*I), it was clearly separated with the second one (*Rsa*I) (Fig. 4). This might be indicative for seasonal changes of either the microbial inoculum or the chemical composition of flood water (Scanlon, 1989), as also suggested by the significant separation of control samples obtained in August 2009.

An example for the contrasting persistence of bacteria with similar (and potentially overlapping) substrate demands is provided by the combined information from FISH and clone libraries from pool NGIII (Figs. 3, 6): Clear differences in the response to extended residence time were found between two groups of bacterial genotypes from lineages that are specialized in the consumption of C1 components. For one, putative methylotrophic β-*Proteobacteria* affiliated with *Methylotenera* were a significant component of all studied pools, as already observed previously for pool NGIII (Shabarova and Pernthaler, 2010). They were introduced in similar quantities by repeated flooding events, and they maintained similar proportions throughout the stagnant periods, as reflected in both, the clone libraries (20 and 30 % of total sequences in NGIII 1 and 3, respectively) and microscopic preparations (Fig. 3). These genotypes are closely related (<4 % difference) to methylotrophic bacteria (*Methylotenera mobilis)* isolated from sediment of lake Washington (Kalyuzhnaya et al.,

2006), and close relatives were detected both, in an aquatic cave environment (Chen et al., 2009) and in the hypolimnetic zone of a freshwater lake (Salcher et al., 2008). By contrast, γ-*proteobacterial* sequence types related to methano- and methylotrophic bacteria from freshwater sediments, drinking water systems and cave water (Stoecker et al., 2006; Rahalkar et al., 2007; Baskar et al., 2012) were only detected shortly after a flooding event (Fig. 6), but not in stagnant pool water. Bacteria from the same lineage, i.e., clustering within the same OTU, were also present in a clone library produced in a similar post-flooding situation during an earlier study (Shabarova and Pernthaler, 2010). This suggests that they represent a recurring component of the microbial assemblages in the karst water of the Bärenschacht system during winter seasons. However, since the closest relative of these genotypes (GU127240, Fig. 5) originates from an anoxic freshwater habitat, it is conceivable that they were passively introduced by flooding from an anoxic or suboxic niche, and were subsequently unable to maintain their populations in the oxygenated pools.

It does not appear entirely appropriate to interpret the observed changes of the microbial assemblages within the conceptual framework of disturbance frequencies and/or intensities (Baho et al., 2012; Berga et al., 2012). For one, the water in the pools was completely exchanged during the flooding events, as indicated by transient water levels that were up to several meters above the normal ones (Fig. 1), leading to a quantitative replacement rather than a disturbance of the bacterioplankton communities. This scenario also somewhat differs from other temporary habitats, e.g., the drying up of ponds or streams (Fazi et al., 2008), because a persistent local seed inoculum is likely much smaller. The studied pools might thus be regarded as rather extreme examples for the importance of mass effects in determining the composition of microbial assemblages (Logue and Lindström, 2008). Secondly, it appears that the diversity of bacteria in the newly introduced karst water was higher than after several weeks of residence in the pools (Fig. 5A), probably because they were collected from a larger variety of niches (Engel, 2010). Thus, in contrast to classical concepts, increasing disturbance frequencies would result in elevated levels of diversity. This was, however, not reflected in our data, possibly due to the limited taxonomic resolution of the T-RFLP approach.

Typical bacteria in karst water during winter season

The majority of large OTUs (encompassing more than two thirds of all sequences) were shared between the microbial assemblages from two winter seasons, and only 6 OTUs ≥3 sequences were not. Ten of 12 β-*proteobacterial* OTUs (affiliated with *Methylophilaceae* and *Comamonadaceae*) were shared both, between the winter seasons and between

samples 1 and 3 from pool NGIII. By contrast, none of the OTUs affiliated with γ-*Proteobacteria* and only 4 of 7 CFB OTUs that were common in both seasons were also shared in several samples within season 2009/2010. This supports our hypotheses that β-*Proteobacteria* represented a core group of karst water bacteria within the studied system, and that changes in the bacterial communities in epiphreatic pools were mainly due to other lineages. In summary, it can be concluded that endokarst communities from the same season harbor large proportions of recurrent populations, which in turn might hint at comparable niche conditions and biogeochemical processes.

Experimental procedures:

Site description, monitoring of physical parameters

Three epiphreatic water pools located in the Northern part of the Bärenschacht cave system (Switzerland) (for a map see Fig. 1 in (Shabarova and Pernthaler, 2010)) were sampled in this study (Fig. 1). All pools are exposed to flood events at different frequencies (Fig. 1). These floodings result from a rise of the water table of the North Sump that is situated in the same cave region and is directly connected to the karst groundwater. Temperature and water level fluctuations in these pools were monitored either with data loggers (Ingenieurbüro Ziegler GmbH), or with an online measuring and reporting system (Cave-Link, Ingenieurbüro Ziegler GmbH) (Fig. 1). The pool NS (average depth 1.5 m), the closest one to the North Sump is also the most frequently flooded one. It forms a separate pool only if the water table of the North Sump is regular, i.e., the pressure sensor next to the North Sump registers values corresponding to a water column of <2.0 m. Otherwise pool NS is linked with the sump. The pool NGIII (average depth 0.4 m) is located at approximately half a kilometer distance from NS in the branch of the 'Galery du Nord'. It is flooded if the karst water table of the North Sump rises higher than 28.5 m above its regular level. Pool LC (average depth 0.2 m) is situated in the 'Longs Couteaux' cave region about one kilometer from NS. Although the altitude of this pool is only a few meters more than that of pool NS, a flood event has to reach at least a magnitude of 41.8 m above the regular level to overcome the geological barrier between the North Sump and the pool.

Sampling procedure

A set of triplicate biomass samples from pool NGIII were obtained on August 23, 2009 (19 days after a flood event), during a pilot campaign that exclusively served for the establishment of the T-RFLP methodology. Sampling for the analysis of the transformations

of the microbial communities in the pools in the context of physicochemical parameters was initiated after the stabilizing of the snow cover (which minimizes the flooding activity) on December 16, 2009 and was continued approximately weekly until January 21, 2010 (6 samples) (Fig. 1). One additional sampling was carried out on March 18, 2010. The sampling campaigns (of durations between 20 and 32 h) were conducted with the technical assistance of experienced speleologists.

At every sampling one liter of water was collected from each system for the determination of chemical parameters (pH, conductivity, alkalinity, concentration of metal ions, silicate, sulfate, phosphate, nitrate and organic carbon parameters). The conductivity, pH and temperature were also measured *in situ* with a portable WTW 340i multimeter, using a SenTix® 41-3 Epoxy pH electrode and an UltraCon® 325 conductivity cell (WTW). Additional 100 ml of water were fixed with 2M HCl to a pH of 2 for additional measurements of dissolved and total organic carbon and nitrogen (DOC, TOC, DN, TN). All samples for chemical analyses were collected in teflon or glass flasks that were prewashed with 2M HCl and heat sterilized. The bottles were rinsed three times with water from the pool where samples were collected. All samples were delivered to the lab within 24 h after sampling.

During each sampling triplicate portions of 50 ml were taken from the pool with sterile syringes for counts of total bacterial abundance and population analysis by catalyzed reporter deposition FISH (CARD-FISH). Samples were fixed immediately with a 37 % formaldehyde solution (FA, 1 % final concentration). At the sampling time points 1, 2, 3, 5 and 7 (Fig. 1) microbial biomass from 1.5 to 2 l of water was collected in duplicates using gamma irradiated Sterivex-GP filter units with polyethersulfone membranes (0.22 μm pore size, Merck Millipore). Subsequently, 1.5 ml of Beads Solution and 0.5 ml of WD1 Solution of the UltraClean® Water DNA isolation kit (MO BIO Laboratories, Inc.) were added to each filter unit and the ends of the unit were sealed for transportation with sterile caps. DNA extraction was conducted within 48 hours after sampling.

Chemical parameters

250 ml of filtrates were prepared in the laboratory within 24 hours after collection. Filtration was conducted through membrane filters (type GTTP, Millipore) that had been pre-washed in 2 M HCl and pre-rinsed with 1 l of deionized sterile water (Milli-Q, Millipore). The filtrates were used for the colorimetric determination of nitrate, nitrite, ammonium, (EDI, 1983), silica concentrations (AA3 Autoanalyzer, Bran&Luebbe). Amount of OC, N (Shimadzu analyzer TOC-V CPH), and of phosphate (AA3 Autoanalyzer, Bran&Luebbe) was quantified for both,unfiltered and filtered samples. The *in situ* measurements of conductivity and pH

were confirmed in the laboratory by means of a portable meter (Titrando 809, Metrohm). An ion chromatograph (Compact IC 761, Metrohm) was used to determine the concentrations of sulphate, chloride and of cations in filtrates.

Cell staining, flow cytometry and microscopic analysis

Flow cytometry (inFlux V-GS, Becton Dickinson) was used to count total bacterial numbers. Triplicates of FA fixed samples (1 ml) were stained with 10 µl of 1/20 stock solution of SYBR® Green I (Invitrogen™). Cells were detected based on their side scatter (SSC) and green fluorescence intensities. Fluorescence was excited using a 488 nm laser (Coherent, Sapphire, 200 mW), and emission was measured at 531 ± 40 nm. Per sample 5 10^4 to 2 10^5 of stained cells were counted. The analysis of the obtained Bi-plots (DNA fluorescence versus SSC) was done with the FlowJo 7.1.2 software (Tree Star, Inc.).

Portions of 7 to 10 ml of FA fixed samples were filtered onto white polycarbonate membrane filters (Millipore GTTP, 0.2 µm pore size; diameter, 47 mm) within 24 h after sampling for analysis by CARD-FISH. Filters were dried and stored at -20 °C. The subsequent embedding of filters, digestion, hybridization and signal amplification were conducted according to previously described protocols (Sekar et al., 2003; Salcher et al., 2008). Horseradish peroxidase labeled probes targeting α, β and γ subgroups of *Proteobacteria*, part of CFB, the *Actinobacteria* (Manz et al., 1992; Roller et al., 1994; Manz et al., 1996; Neef, 1997) and the *Methylophilaceae* (Friedrich et al., 2003) were used for analysis of bacterial community composition. After CARD-FISH, the filters were counterstained with DAPI and the relative abundance of hybridized cells in 500–1000 DAPI-stained cells were counted manually by epifluorescence microscopy (AxioImager Z1, Carl Zeiss) at UV and blue (Zeiss filter sets 01, 09) excitation.

DNA extraction and T-RFLP analysis

DNA was extracted with the UltraClean® Water DNA isolation kit (MO BIO Laboratories, Inc.) under sterile conditions. The membranes and the added liquid content (see sampling procedure) of the Sterivex filter units were transferred to tubes containing beads from the DNA isolation kit. Beads solution amounting to 2.5 ml was added and the tube was vortexed for 10 min. DNA extraction steps were conducted according to the manufacturer's instructions. DNA was eluted with 2 ml of UltraPure™ water (Invitrogen™) and frozen at -20°C until future processing.

For T-RFLP analysis 16S rRNA genes were amplified using primers GM3F (FAM labeled) and GM4R (Muyzer et al., 1995). PCR was performed using Platinum® PCR SuperMix High Fidelity (Invitrogen™). Amplification included denaturation for 5 min at 94 °C, and 26 subsequent cycles with denaturation for 30 s at 94 °C, annealing for 30 s at 48 °C, and extension for 90 s at 68 °C, followed by a final extension for 10 min at 68 °C and cooling to 10 °C. The quality of PCR products was examined by electrophoresis in 1 % w/v agarose gel with subsequent GelRed™ (Biotium, Inc.) staining. The quantity of DNA was estimated after purification of the amplification products (QIAquick PCR purification kit, QIAGEN) with the help of a NanoPhotometer™ (Implen GmbH). Purified PCR product (300 ng) were digested overnight at 37 °C with 6 U of the restriction endonucleases *Msp*I or *Rsa*I (Promega) in 20 µl of restriction enzyme buffers (Promega). After addition of GeneScan™-1000 ROX™ size standard (Applied Biosystems) digestion products were analyzed using an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) as previously described (Hartmann et al., 2005; Hartmann and Widmer, 2008). The amount of digested DNA per sample was optimized to obtain maximal resolution without overloading of the capillaries.

GeneMarker® version 1.51 (SoftGenetics LCC) was used for analysis of T-RF peak sizes and heights in the obtained electropherograms. Only fragments longer than 50 base pairs were used for the subsequent analysis of terminal fragment length polymorphisms of amplification products. Peaks below a threshold level of 50 relative fluorescence units (rfu) and only appearing in a single sample were excluded. For data normalization individual peak height values were divided by the sum of all peak height values in the corresponding profile to correct for differences in sample loading. The phylogenetic affiliation of the detected fragments was predicted using a sequence dataset from a clone library of the NGIII system that was obtained with the same primers in winter 2008-2009 (Shabarova and Pernthaler, 2010). Agglomerative hierarchic clustering was performed for complete profiles and after exclusion of particular bacterial groups, as identified by the *in silico* prediction. The software SPSS Statistics 20 (IBM) and XLSTAT 11 (Addinsoft) were used to construct dendograms calculated according to the *Pearson correlation coefficient (complete linkage) method, and to perform the Mantel test.*

16S rRNA gene clone libraries

DNA obtained from samples NGIII 1 and NGIII 3 was used for the creation of clone libraries. For this purpose, 16S rRNA genes were amplified using primers GM3F and GM4R as described above for T-RFLP analysis. The purified PCR products were cloned using the pGEM®-T Vector System I cloning kit (Promega). GenElute™ Five-Minute Plasmid Miniprep Kit (Sigma-Aldrich GmbH) was used for purification of insert-bearing plasmids. Screening

was conducted by PCR using M13 vector primers and GoTaq® Green Master Mix (Promega). Subsequent sequencing of 16S rRNA genes from plasmid inserts was performed on an ABI 3130x Genetic Analyzer using the ABI BigDye chemistry (Applied Biosystems) and primers GM1F (Muyzer et al., 1993) and GM3F. Partial sequences were assembled using ContigExpress of Vector NTI package (Invitrogen™). Chimeric sequences were identified and removed with help of the Mallard and Pintail software (Ashelford et al., 2005). Sequences were submitted to the EMB database under the accessions HE998777-HE999442.

Phylogenetic and diversity analyses

The SINA web aligner tool of the SILVA project (Pruesse et al., 2012) was used for alignment of 16S rRNA gene sequences. The aligned sequences were added to the reference tree of the SILVA 16S rRNA database (release 108) using the 'Parsimony_quick-add' function of ARB release 5.3 (Ludwig, *et al*., 2004). The alignments of phylogenetic clusters of interest were subsequently manually optimized and the positions of sequences were recalculated. The 'Parsimony_interactive' tool of ARB was then used for the optimization of tree topologies of *Bacteroidetes* and *Methylococaceae*. Only sequences >1200 bases and Pintail values >90 were used for this purpose. Subsequently, maximum likelihood (ML) analyses were carried out for *Methylococaceae* to establish the closest phylogenetic neighbors. Finally, up to 550 of close related sequences were used for the construction of bootstrapped ML trees (1000 repetitions) applying the General Time Reversible CAT approximation of the RAxML software (v7.2.8 alpha release) (Stamatakis, 2006a; Stamatakis, 2006b). The resulting trees were imported into ARB and nodes with bootstrap values <50 % were collapsed into multifurcations.

The number of OTUs in the clone libraries, their overlap and the development of predicted diversity estimates (Chao 1) in differently sized subsamples were calculated at a similarity level of 99 % using the DOTUR and SONS software (Schloss and Handelsman, 2005; Schloss and Handelsman, 2006) and a dedicated online tool (Kemp and Aller, 2004) as previously described (Shabarova and Pernthaler, 2010). The closest relative of sequences within selected OTUs of *Comamonadaceae* was identified by analysis of a distance matrix obtained from their pairwise alignment with all sequences in the SILVA database. For the identification of typical karst water bacteria, information from two more clone libraries from pool NGIII was included, one from a sampling in winter 2007/2008 (102 sequences)(Shabarova and Pernthaler, 2010), and one from the last sampling date (62 sequences).

Acknowledgements:

We would like to thank the cavers of SGH Bern especially Diego Sanz, Daniel Burkhalter, Martin Achtman and Stefan Huber for their help in preparing and conducting the sampling campaigns; Hans-Peter Kohler and the “Labor Analytik und Ausbildung” at the Swiss Federal Institute of Aquatic Science & Technology for chemical analyses; Salome Schneider and Roy Kissling for the introduction to T-RFLP; Jörg Villiger for the automated identification of relatives of *Comamonadaceae*; and Ester Eckert for discussions and backup. T.S was supported by a stipend of the Research Commission of the University of Zurich (grant Nr. 57150401).

Legends to Figures:

Figure 1: Timing and magnitude of flooding events of the studied epiphreatic pools in the Bärenschacht cave system. The left half of the diagram schematically illustrates the hydrological connection between pools and the required rise in the water table for a flooding event to occur. Black symbols, location of temperature/pressure sensors; dashed lines: time points of samplings. Inset (A): magnified depiction of flooding event that occurred in pool NGIII between samplings 3 and 4.

Figure 2: Changes of concentrations of dissolved organic carbon (DOC) and conductivity in the three epiphreatic pools. A break in the dashed line indicates the occurrence of flood events (represented by triangles above panels). White and grey background colours represent water originating from different floods.

Figure 3: Changes in the numbers of β-*Proteobacteria* and of *Methylophilaceae* in the three epiphreatic pools. A break in the dashed line indicates the occurrence of flood events (represented by triangles above panels). White and grey background colours represent water originating from different floods.

Figure 4: Similarity of T-RFLP patterns of the microbial assemblages in the three epiphreatic pools at different time points and after removal of β-*proteobacterial* T-RF peaks as predicted by *in silico* analysis. Left and right panels depict analysis using restriction enzymes *Rsa*I and *Msp*I, respectively. Data labelled NGIII-0 A, B, C refer to a set of triplicate water samples collected from pool NGIII in August 2009. Differences above the dotted line are significant at $P<0.05$. Numbers in brackets: days since last flooding.

Figure 5: Diversity analysis of clone libraries from pool NGIII four days after flooding (NGIII 1) and two weeks later (NGIII 3). Predicted numbers of (A) all bacterial and (B) β-*proteobacterial* OTUs (Chao 1) estimated for different library sizes. Inset panels in (A) and (B): rarefaction curves corresponding to each data set. (C) Microdiversity analysis of β-*proteobacterial* OTUs (defined at 99 % similarity) affiliated with *Comamonadaceae*: number of sequences that have a common closest neighbour in the SILVA database (Release 106). The numbers above the bars give the distance to the closest neighbour (%); broken lines separate individual OTUs. Bar colours: white, black: only present in NGIII 1 or NGIII 3 respectively; grey: present in both samples

Figure 6: Phylogenetic relatedness of *Methylococcaceae* from pool NGIII, as reconstructed by Maximum Likelihood analysis applying the General Time Reversible CAT approximation of the RAxML. Node values are bootstrap values of 1000 replicate samplings. Nodes with values <50 % were resolved into multifurcations

559 **Supplementary material:**

560 **Supplementary Figure S1:** Changes in total bacterial numbers in the three epiphreatic
561 pools. A break in the dashed line indicates the occurrence of flood events (represented by
562 triangles above panels). White and grey background colours represent water originating from
563 different floods.

564 **Supplementary Figure S2:** Phylogenetic affiliation of *Cytophagaceae* and
565 *Flavobacteriaceae* in pool NGIII. The colour gradient symbolized the respective proportions
566 of sequences from the two libraries within individual clades.

References

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J., and Weightman, A.J. (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology* **71**: 7724-7736.

Baho, D.L., Peter, H., and Tranvik, L.J. (2012) Resistance and resilience of microbial communities - temporal and spatial insurance against perturbations. *Environ Microbiol* **19**: 1462-2920.

Barton, H.A., and Luiszer, F. (2005) Microbial metabolic structure in a sulfidic cave hot spring: Potential mechanisms of biospeleogenesis. *Journal of Cave and Karst Studies* **67**: 28-38.

Baskar, S., Baskar, R., Thorseth, I.H., Ovreas, L., and Pedersen, R.B. (2012) Microbially Induced Iron Precipitation Associated with a Neutrophilic Spring at Borra Caves, Vishakhapatnam, India. *Astrobiology* **12**: 327-346.

Berga, M., Szekely, A.J., and Langenheder, S. (2012) Effects of disturbance intensity and frequency on bacterial community composition and function. *PLoS One* **7**: 14.

Butscher, C., Auckenthaler, A., Scheidler, S., and Huggenberger, P. (2011) Validation of a Numerical Indicator of Microbial Contamination for Karst Springs. *Ground Water* **49**: 66-76.

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H. et al. (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J* **3**: 1093-1104.

EDI (1983) *Richtlinien für die Untersuchung von Abwasser und Oberflächenwasser, Allgemeine Hinweise und Analysenmethoden*: Eidgenössisches Departement des Innern (EDI).

Engel, A.S. (2010) Microbial diversity of cave ecosystems. In *Geomicrobiology: Molecular and Environmental Perspectives*. Barton, L., Mandl, M., and Loy, A. (ed): Springer, pp. 219-238.

Farnleitner, A.H., Wilhartitz, I., Ryzinska, G., Kirschner, A.K., Stadler, H., Burtscher, M.M. et al. (2005) Bacterial dynamics in spring water of alpine karst aquifers indicates the presence of stable autochthonous microbial endokarst communities. *Environ Microbiol* **7**: 1248–1259.

Fazi, S., Amalfitano, S., Piccini, C., Zoppini, A., Puddu, A., and Pernthaler, J. (2008) Colonization of overlaying water by bacteria from dry river sediments. *Environmental Microbiology* **10**: 2760-2772.

Ford, D.C., and Williams, P. (2007) *Karst Hydrogeology and Geomorphology*: Wiley.

Friedrich, U., Van, L., H, Altendorf, K., and Lipski, A. (2003) Microbial community and physicochemical analysis of an industrial waste gas biofilter and design of 16S rRNA-targeting oligonucleotide probes. *Environ Microbiol* **5**: 183–201.

Gams, I. (2003) *Kras v Sloveniji v prostoru i času (Karst in Slovenia in space and time).* Ljubljana: Založba ZRC, ZRC SAZU.

Griebler, C., and Lueders, T. (2009) Microbial biodiversity in groundwater ecosystems. *Freshwater Biology* **54**: 649-677.

607 Hahn, M., and Pockl, M. (2005) Ecotypes of planktonic Actinobacteria with identical 16S
608 rRNA genes adapted to thermal niches in temperate, subtropical, and tropical freshwater
609 habitats. *Applied and Environmental Microbiology* **71**: 766-773.

610 Hahn, M.W., Kasalicky, V., Jezbera, J., Brandt, U., Jezberova, J., and Simek, K. (2010)
611 Limnohabitans curvus gen. nov., sp nov., a planktonic bacterium isolated from a freshwater
612 lake. *International Journal of Systematic and Evolutionary Microbiology* **60**: 1358-1365.

613 Hartmann, M., and Widmer, F. (2008) Reliability for detecting composition and changes of
614 microbial communities by T-RFLP genetic profiling. *FEMS Microbiol Ecol* **63**: 249-260.

615 Hartmann, M., Frey, B., Kolliker, R., and Widmer, F. (2005) Semi-automated genetic
616 analyses of soil microbial communities: comparison of T-RFLP and RISA based on
617 descriptive and discriminative statistical approaches. *J Microbiol Methods* **61**: 349-360.

618 Häuselmann, P. (2002) Cave genesis and its relationship to surface processes:
619 Investigations in the Siebenhengste region (BE, Switzerland). In *Höhlenforschung im Gebiet*
620 *Sieben Hengste-Hohgant*. Fribourg: University of Fribourg.

621 Hoshino, T., Yilmaz, L.S., Noguera, D.R., Daims, H., and Wagner, M. (2008) Quantification
622 of target molecules needed to by fluorescence in situ hybridization (FISH) and catalyzed
623 reporter deposition-FISH. *Applied and Environmental Microbiology* **74**: 5068-5077.

624 Iker, B.C., Kambesis, P., Oehrle, S.A., Groves, C., and Barton, H.A. (2010) Microbial atrazine
625 breakdown in a karst groundwater system and its effect on ecosystem energetics. *J Environ*
626 *Qual* **39**: 509-518.

627 Jaspers, E., Nauhaus, K., Cypionka, H., and Overmann, J. (2001) Multitude and temporal
628 variability of ecological niches as indicated by the diversity of cultivated bacterioplankton.
629 *FEMS Microbiol Ecol* **36**: 153–164.

630 Jezbera, J., Jezberova, J., Brandt, U., and Hahn, M.W. (2011) Ubiquity of Polynucleobacter
631 necessarius subspecies asymbioticus results from ecological diversification. *Environmental*
632 *Microbiology* **13**: 922-931.

633 Kalyuzhnaya, M.G., Bowerman, S., Lara, J.C., Lidstrom, M.E., and Chistoserdova, L. (2006)
634 Methylotenera mobilis gen. nov., sp. nov., an obligately methylamine-utilizing bacterium
635 within the family Methylophilaceae. *Int J Syst Evol Microbiol* **56**: 2819–2823.

636 Kemp, P.F., and Aller, J.Y. (2004) Estimating prokaryotic diversity: When are 16S rDNA
637 libraries large enough? *Limnology and Oceanography: Methods* **2**: 114–125.

638 Logue, J.B., and Lindström, E.S. (2008) Biogeography of bacterioplankton in inland waters.
639 *Freshwater Reviews* **1**: 99-114.

640 Macalady, J.L., Dattagupta, S., Schaperdoth, I., Jones, D.S., Druschel, G.K., and Eastman,
641 D. (2008) Niche differentiation among sulfur-oxidizing bacterial populations in cave waters.
642 *ISME J* **2**: 590–601.

643 Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K.-H. (1992) Phylogenetic
644 oligodeoxynucleotide probes for the major subclasses of *Proteobacteria*: Problems and
645 solutions. *Systematic and Applied Microbiology* **15**: 593-600.

646 Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.-H. (1996) Application of
647 a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the

phylum *Cytophaga-Flavobacter-Bacteroides* in the natural environment. *Microbiology* **142**: 1097-1106.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.

Muyzer, G., Teske, A., Wirsen, C., and Jannasch, H. (1995) Phylogenetic relationship of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gel electrophoresis of 16S rDNA fragments. *Arch Microbiol* **164**: 165–172.

Neef, A. (1997) Anwendung der *in situ*-Einzelzell-Identifizierung von Bakterien zur Populationsanalyse in komplexen mikrobiellen Biozönosen. In *Lehrstuhl für Mikrobiologie*. München: Techical University, Munich, Germany.

Pronk, M., Goldscheider, N., and Zopfi, J. (2006) Dynamics and interaction of organic carbon, turbidity and bacteria in a karst aquifer system. *Hydrogeology Journal* **14**: 473-484.

Pronk, M., Goldscheider, N., and Zopfi, J. (2007) Particle-size distribution as indicator for fecal bacteria contamination of drinking water from karst springs. *Environ Sci Technol* **41**: 8400–8405.

Pronk, M., Goldscheider, N., and Zopfi, J. (2009) Microbial communities in karst groundwater and their potential use for biomonitoring. *Hydrogeology Journal* **17**: 37–48.

Pruesse, E., Peplies, J., and Glöckner, F.O. (2012) SINA: accurate high throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*.

Rahalkar, M., Bussmann, I., and Schink, B. (2007) Methylosoma difficile gen. nov., sp nov., a novel methanotroph enriched by gradient cultivation from littoral sediment of Lake Constance. *International Journal of Systematic and Evolutionary Microbiology* **57**: 1073-1080.

Ravbar, N., Engelhardt, I., and Goldscheider, N. (2011) Anomalous behaviour of specific electrical conductivity at a karst spring induced by variable catchment boundaries: the case of the Podstenjsek spring, Slovenia. *Hydrological Processes* **25**: 2130-2140.

Reischer, G.H., Haider, J.M., Sommer, R., Stadler, H., Keiblinger, K.M., Hornek, R. et al. (2008) Quantitative microbial faecal source tracking with sampling guided by hydrological catchment dynamics. *Environmental Microbiology* **10**: 2598-2608.

Roller, C., Wagner, M., Amann, R., Ludwig, W., and Schleifer, K.-H. (1994) *In situ* probing of Gram-positive bacteria with high DNA G+C content using 23S rRNA-targeted oligonucleotides. *Microbiology* **140**: 2849-2858.

Salcher, M.M., Pernthaler, J., Zeder, M., Psenner, R., and Posch, T. (2008) Spatio-temporal niche separation of planktonic Betaproteobacteria in an oligo-mesotrophic lake. *Environ Microbiol* **10**: 2074–2086.

Scanlon, B.R. (1989) Physical Controls on Hydrochemical Variability in the Inner Bluegrass Karst Region of Central Kentuckya. *Ground Water* **27**: 639-646.

Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology* **71**: 1501–1506.

689 Schloss, P.D., and Handelsman, J. (2006) Introducing SONS, a tool for operational
690 taxonomic unit-based comparisons of microbial community memberships and structures.
691 *Applied and Environmental Microbiology* **72**: 6773–6779.

692 Sekar, R., Pernthaler, A., Pernthaler, J., Warnecke, F., Posch, T., and Amann, R. (2003) An
693 improved protocol for quantification of freshwater Actinobacteria by fluorescence in situ
694 hybridization. *Applied and Environmental Microbiology* **69**: 2928-2935.

695 Shabarova, T., and Pernthaler, J. (2010) Karst pools in subsurface environments: collectors
696 of microbial diversity or temporary residence between habitat types. *Environmental*
697 *Microbiology* **12**: 1061-1074.

698 Simon, K., Pipan, T., and Culver, D. (2007) A conceptual model of the flow and distribution of
699 organic carbon in caves. *Journal of Cave and Karst Studies* **69**: 279–284.

700 Simon, K.S., Pipan, T., Ohno, T., and Culver, D.C. (2010) Spatial and temporal patterns in
701 abundance and character of dissolved organic matter in two karst aquifers. *Fundamental and*
702 *Applied Limnology* **177**: 81-92.

703 Stamatakis, A. (2006a) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses
704 with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688-2690.

705 Stamatakis, A. (2006b) Phylogenetic models of rate heterogeneity: a high performance
706 computing perspective. In *Proceedings of the 20th international conference on Parallel and*
707 *distributed processing*. Rhodes Island, Greece: IEEE Computer Society, pp. 253-253.

708 Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012). Stochastic and
709 deterministic assembly processes in subsurface microbial communities [WWW document].
710 URL http://dx.doi.org/10.1038/ismej.2012.22.

711 Stoecker, K., Bendinger, B., Schoning, B., Nielsen, P.H., Nielsen, J.L., Baranyi, C. et al.
712 (2006) Cohn's Crenothrix is a filamentous methane oxidizer with an unusual methane
713 monooxygenase. *Proceedings of the National Academy of Sciences of the United States of*
714 *America* **103**: 2363-2367.

715 Wilhartitz, I.C., Kirschner, A.K., Stadler, H., Herndl, G.J., Dietzel, M., Latal, C. et al. (2009)
716 Heterotrophic prokaryotic production in ultraoligotrophic alpine karst aquifers and ecological
717 implications. *FEMS Microbiol Ecol* **68**: 287-299.

Figure 1

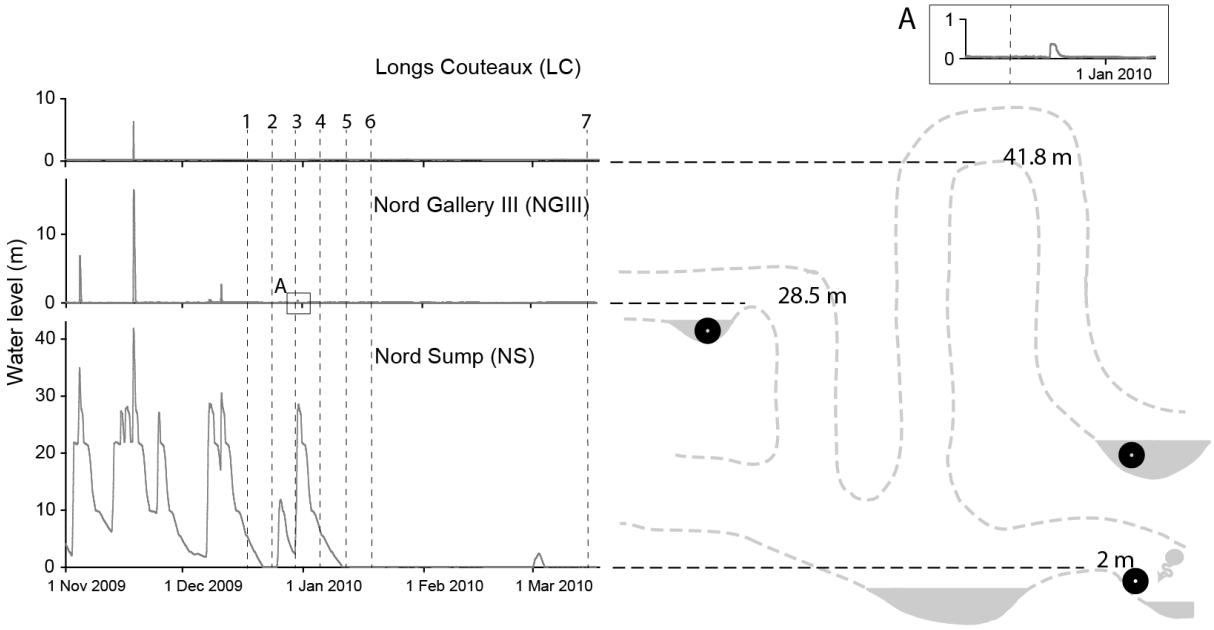

A
1
0
1 Jan 2010
Longs Couteaux (LC)
10
0
1 2 3 4 5 6 7
41.8 m
Nord Gallery III (NGIII)
10
0
A
28.5 m
Water level (m)
40
30
20
10
0
Nord Sump (NS)
2 m
1 Nov 2009
1 Dec 2009
1 Jan 2010
1 Feb 2010
1 Mar 2010

Figure 2

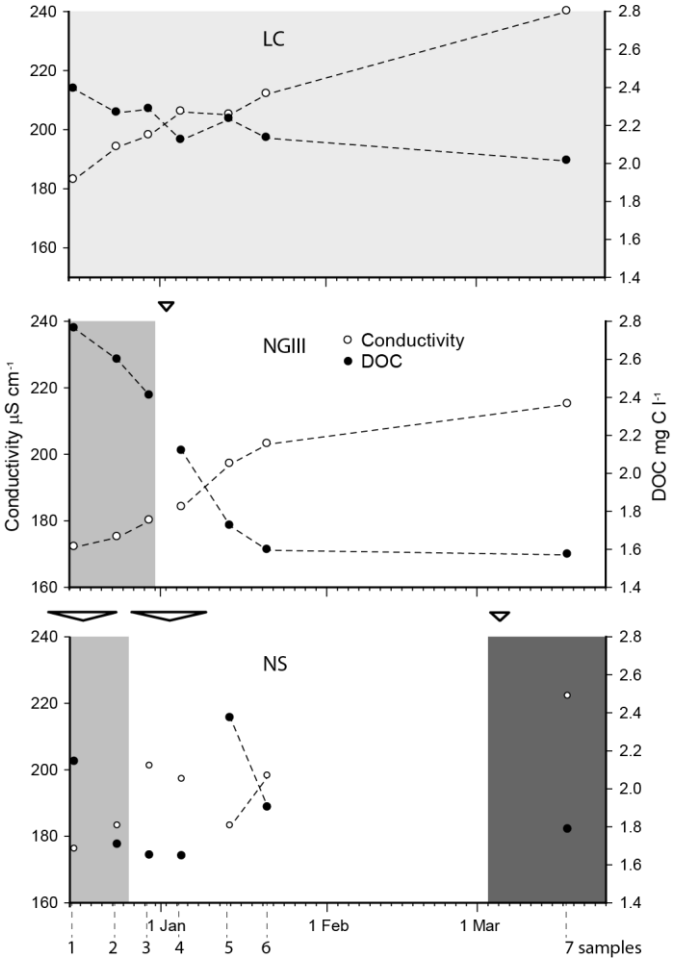

LC
NGIII
NS
Conductivity
DOC
Conductivity µS cm-1
DOC mg C l-1
240
220
200
180
160
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1 Jan
1 Feb
1 Mar
1
2
3
4
5
6
7 samples

Figure 3

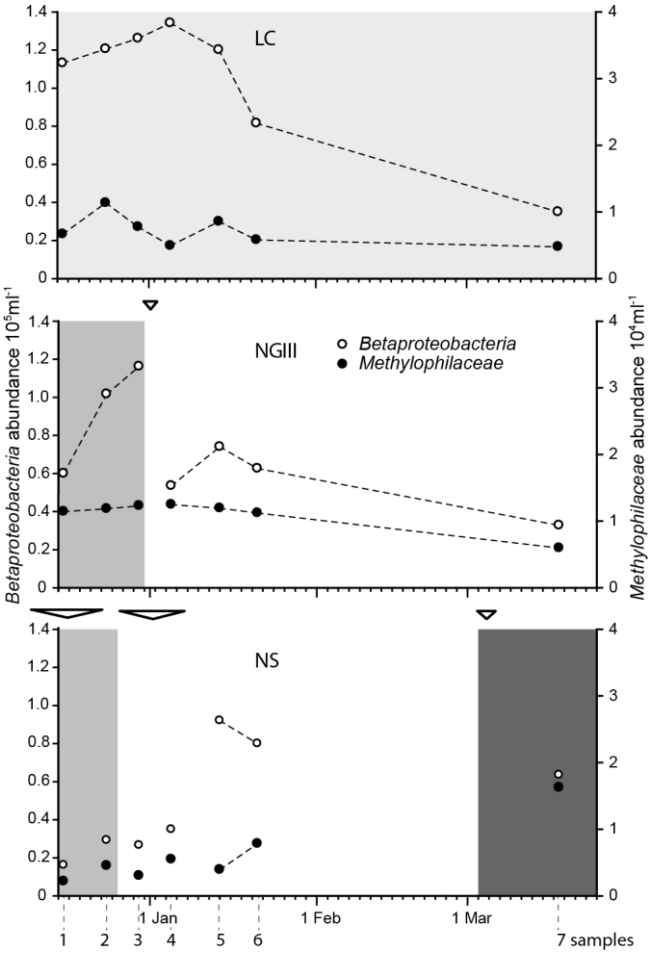

LC
NGIII
NS
Betaproteobacteria
Methylophilaceae
Betaproteobacteria abundance 10^5 ml^-1
Methylophilaceae abundance 10^4 ml^-1
1 Jan
1 Feb
1 Mar
1
2
3
4
5
6
7 samples

Figure 4

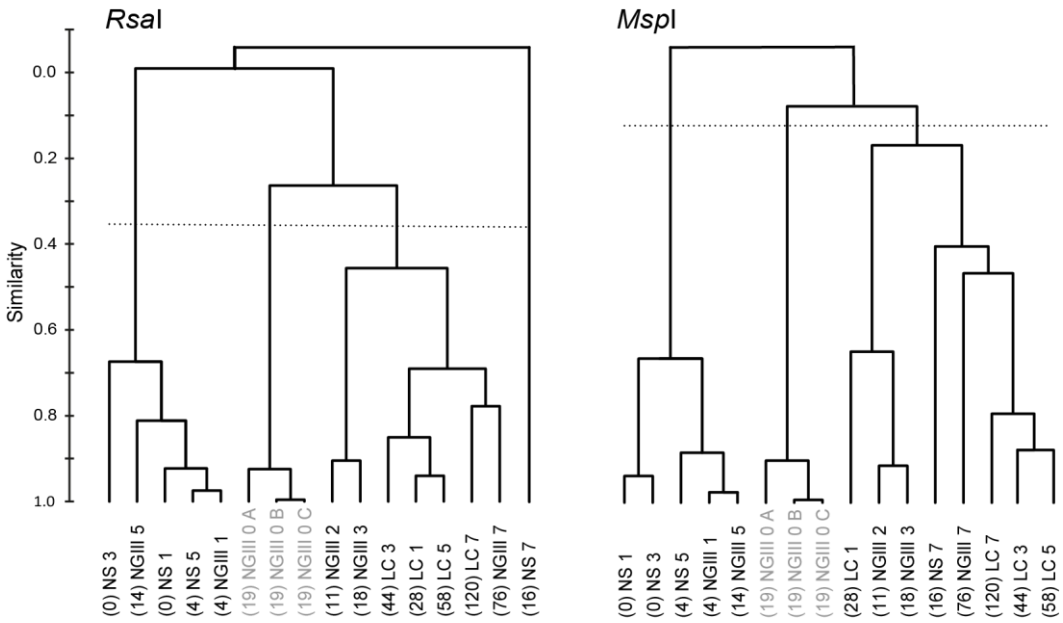

RsaI
MspI
Similarity
0.0
0.2
0.4
0.6
0.8
1.0
(0) NS 3
(14) NGIII 5
(0) NS 1
(4) NS 5
(4) NGIII 1
(19) NGIII 0 A
(19) NGIII 0 B
(19) NGIII 0 C
(11) NGIII 2
(18) NGIII 3
(44) LC 3
(28) LC 1
(58) LC 5
(120) LC 7
(76) NGIII 7
(16) NS 7
(0) NS 1
(0) NS 3
(4) NS 5
(4) NGIII 1
(14) NGIII 5
(19) NGIII 0 A
(19) NGIII 0 B
(19) NGIII 0 C
(28) LC 1
(11) NGIII 2
(18) NGIII 3
(16) NS 7
(76) NGIII 7
(120) LC 7
(44) LC 3
(58) LC 5

Figure 5

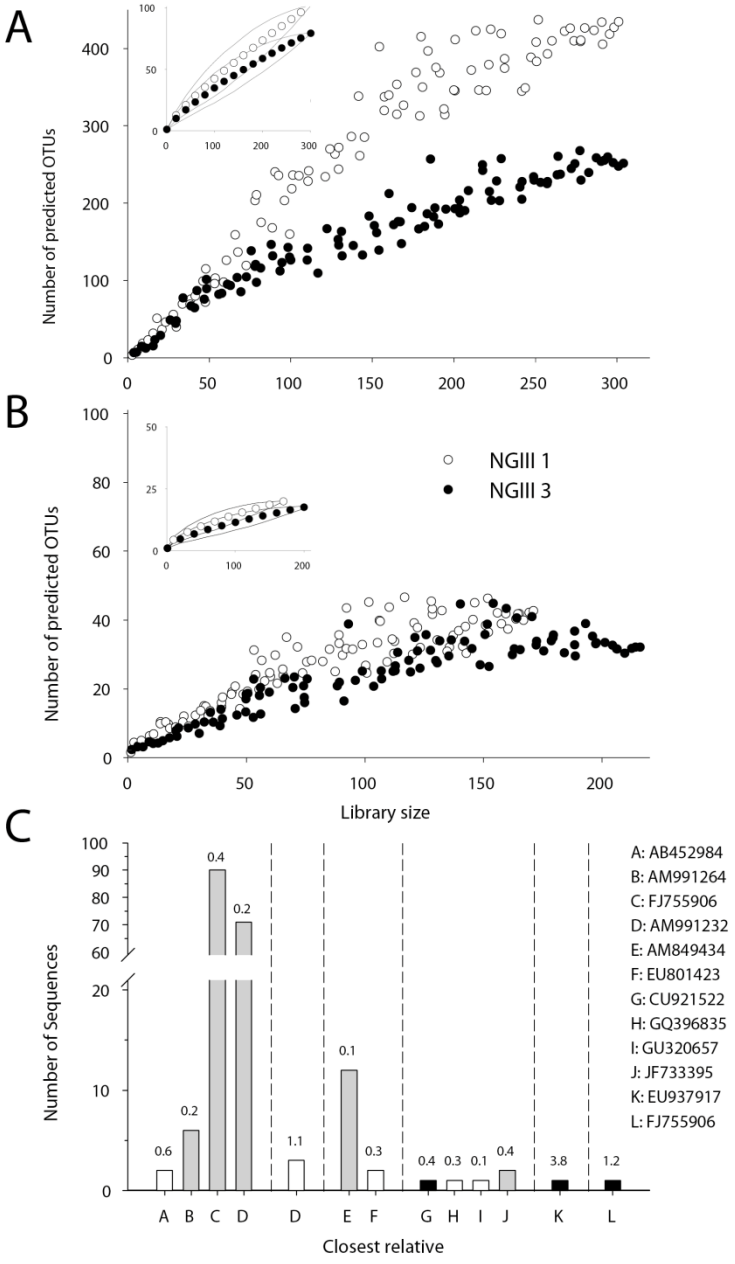
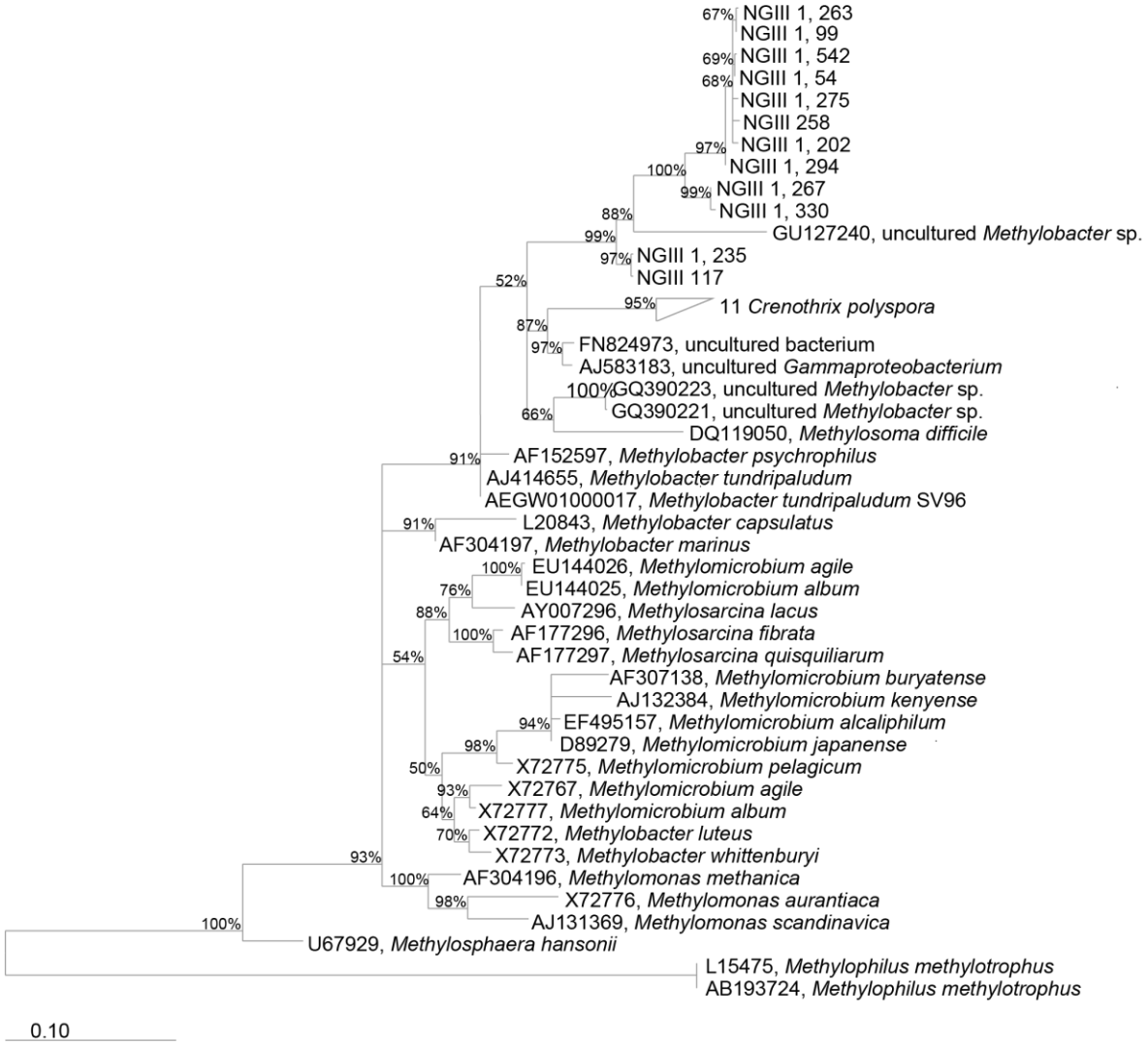
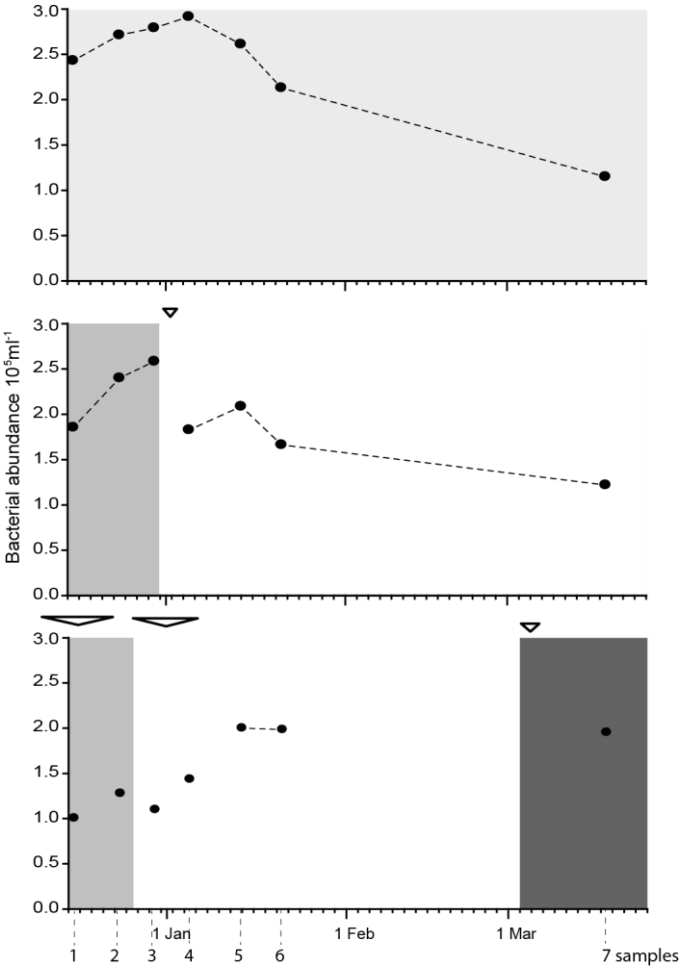

A
Number of predicted OTUs
B
Number of predicted OTUs
NGIII 1
NGIII 3
Library size
C
Number of Sequences
A: AB452984
B: AM991264
C: FJ755906
D: AM991232
E: AM849434
F: EU801423
G: CU921522
H: GQ396835
I: GU320657
J: JF733395
K: EU937917
L: FJ755906
Closest relative

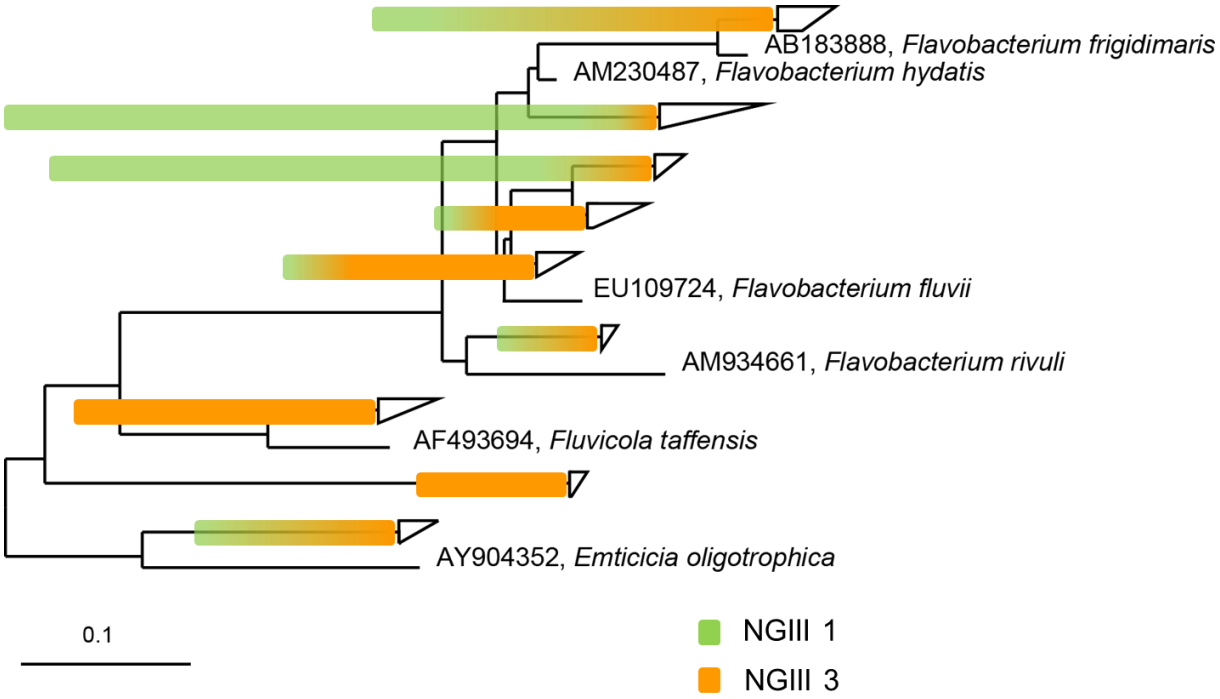
Figure 6


67%
NGIII 1, 263
NGIII 1, 99
69%
NGIII 1, 542
68%
NGIII 1, 54
NGIII 1, 275
NGIII 258
97%
NGIII 1, 202
100%
NGIII 1, 294
99%
NGIII 1, 267
NGIII 1, 330
88%
GU127240, uncultured Methylobacter sp.
99%
97%
NGIII 1, 235
NGIII 117
52%
95%
11 Crenothrix polyspora
87%
97%
FN824973, uncultured bacterium
AJ583183, uncultured Gammaproteobacterium
100%
GQ390223, uncultured Methylobacter sp.
66%
GQ390221, uncultured Methylobacter sp.
DQ119050, Methylosoma difficile
91%
AF152597, Methylobacter psychrophilus
AJ414655, Methylobacter tundripaludum
AEGW01000017, Methylobacter tundripaludum SV96
91%
L20843, Methylobacter capsulatus
AF304197, Methylobacter marinus
100%
EU144026, Methylomicrobium agile
76%
EU144025, Methylomicrobium album
88%
AY007296, Methylosarcina lacus
100%
AF177296, Methylosarcina fibrata
54%
AF177297, Methylosarcina quisquiliarum
AF307138, Methylomicrobium buryatense
AJ132384, Methylomicrobium kenyense
94%
EF495157, Methylomicrobium alcaliphilum
98%
D89279, Methylomicrobium japanense
50%
X72775, Methylomicrobium pelagicum
93%
X72767, Methylomicrobium agile
64%
X72777, Methylomicrobium album
70%
X72772, Methylobacter luteus
X72773, Methylobacter whittenburyi
93%
100%
AF304196, Methylomonas methanica
98%
X72776, Methylomonas aurantiaca
AJ131369, Methylomonas scandinavica
100%
U67929, Methylosphaera hansonii
L15475, Methylophilus methylotrophus
AB193724, Methylophilus methylotrophus
0.10

Supplementary Figure S1



Supplementary Figure S2



Manuscript IV

Bacterial community structure and dissolved organic matter in subsurface karst water pools: interplay between habitat filtering and flooding events

Submitted to the ISME Journal

Bacterial community structure and dissolved organic matter in subsurface karst water pools: interplay between habitat filtering and flooding events

Tanja Shabarova[1], Jörg Villiger[1], Oleg Morenkov[3], Jutta Niggemann[2], Thorsten Dittmar[2] & Jakob Pernthaler[1]*

[1]*Limnological Station, Institute of Plant Biology, University of Zurich, Seestrasse 187, CH-8802 Kilchberg, Switzerland*

[2]*Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl-von-Ossietzky-Strasse, 9-11, D-26129 Oldenburg, Germany*

[3]*Hewlett-Packard Company, Schwarzenburgstrasse 160, CH-3097 Liebefeld, Switzerland*

Running head: microbes in subsurface karst water pools

*corresponding author: Phone: 0041 634 92 10; Fax: 0041 634 92 25; Email: pernthaler@limnol.uzh.ch

Abstract

Bacterial diversity, community structure, and the composition of the dissolved organic carbon (DOC) were studied by Next Generation Sequencing and high resolution mass spectrometry in three epiphreatic subsurface karst water pools with different flooding regimes. We hypothesized that microbes collected from various subsurface niches would face environmental filtering in the pools towards a 'typical' karst water community, and that this selection would be reflected in specific transformations of DOC. Analysis of phylogenetic clustering revealed that deterministic processes had already shaped the bacterial assemblages of the influx water. While these initial communities harbored lineages with large distances to the known diversity, e.g., from the candidate division OD1, the 'core' microbiota of the pools (i.e., present in all samples) were predominantly related to bacteria from subsurface or surface aquatic environments. As predicted, the stagnation of karst water in the pools reproducibly resulted in a decline of bacterial diversity and a community shift towards a subset of closely related genotypes. A small number of operational taxonomic units (OTUs) related to *Betaproteobacteria* were responsible for a significant clustering of samples according to water residence time, and a strong temporal signal was also found in OTUs from several other lineages. By contrast, several OTUs affiliated with *Cytophagaceae* and *Flavobacteriaceae* were more tightly related to different flooding events. These floodings were also a main determinant of DOC composition and could be separated by a small set of molecular formulae with discernible chemical properties. In parallel, a distinct transformation of DOC quality was observed over time. This suggests that the 'core' microbial assemblages in the pools were well-adapted to utilize the available DOC sources.

Introduction

High-throughput ‘Next Generation’ sequence analysis (NGS) of rRNA genes has promoted our understanding of natural microbial assemblages. The comprehensive coverage of microbial rRNA sequences from a particular habitat provides a more precise estimate of total genotype richness, as well as of the respective shares of already ‘known’ diversity (i.e., of genotypes that are closely related to entries in public databases) and of unknown lineages, e.g., of the ‘rare biosphere’ (Sogin et al 2006). Moreover, NGS also represents a tool for testing hypotheses related to (meta)community ecology, as it allows for the application of powerful statistical techniques to compare phylogenetic community structure and turnover, e.g., in fragmented or interconnected habitats (Langenheder et al 2012, Stegen et al 2012, Webb et al 2002).

Microbial communities in temporary aquatic habitats are shaped by the interplay of different assembly processes that may, e.g., be related to temporal changes of abiotic factors (Langenheder et al 2012). On the one hand, their composition may to a large extent be shaped by microbes introduced by the influx. Depending on ecotone connectivity, these inocula may originate from a single or from various habitats (Fazi et al 2008), and it is conceivable that they are, moreover, transformed by the transport process itself. On the other hand, local selective processes may act upon microbial assemblages during their residence in a temporary habitat, e.g., related to quality and quantity of the dissolved organic carbon (DOC) (Kirchman et al 2004).

Organic matter in aquatic habitats is a complex mixture of compounds that are more or less accessible to microbial degradation (Amon and Benner 1994), and that, in turn, is transformed by microbial metabolism (Berggren et al 2009). Novel analytical techniques such as Fourier-Transform Ion Cyclotron Resonance-Mass Spectrometry (FT-IRC-MS) can provide detailed information about DOC composition by quantitative separation of large sets of molecular formulae (Koch et al 2005, Kujawinski et al 2002). This allows the distinction of DOC of different origin, e.g., of terrestrial/freshwater influx into marine waters (Dittmar et al 2006) or to investigate its degradation within a habitat (Gonsior et al 2009). A multitude of factors may influence the changes of DOC composition in surface water habitats, e.g., allochthonous influx (Berggren, et al., 2009) and light-dependent processes such as primary production (Hama and Yanagi 2001) and photodegradation (Bertilsson and Tranvik 2000). Thus, the role of microbes in DOC transformation might be most easily apprehended in natural aquatic systems that are temporarily stagnant and not exposed to light.

The subsurface aquatic karst environment is important with respect to water resource protection and management (Ford and Williams 2007). It also represents a landscape of

exclusive microbial habitats (Engel 2010b), some of which are readily accessible. While there is information about microbes in vadose (unsaturated) streams (Simon et al 2010) and phreatic (saturated) aquifers (Zhou et al 2012), the epiphreatic (periodically flooded) realms of karst are largely unexplored. Temporary pools in the epiphreatic zone harbor highly diverse microbial communities that are collected from different karstic and karst-related ecosystems during flooding events, and that transform during interjacent periods of stability (Shabarova et al 2013). Since both, the inocula to neighboring pools and the local physicochemical conditions are highly similar, these systems are interesting models to assess the role of deterministic and stochastic processes for the development of planktonic bacterial assemblages in the context of carbon limitation.

We evaluated the diversity changes of the microbial assemblages in three epiphreatic pools of a karst cave system at different time points by NGS. Specifically, we investigated the existence of a ‘core’ karst water microbiome, the processes governing microbial community assemblage, and we tested the hypothesis that microbial diversity would reproducibly decrease in pools during periods of stability (Shabarova et al 2013). In parallel we analyzed the DOC composition in these pools by FT-IRC-MS to distinguish between flood-specific molecular indicators and chemical signals of microbial DOC-transformation.

Experimental procedures

Sampling procedure

Three epiphreatic pools of the Bärenschacht cave system (Switzerland), termed 'Longs Couteaux' (LC), 'Galery du Nord III' (NGIII), and North Sump (NS) exposed to repeated karst water flooding events were sampled during seven expeditions between Dec 16, 2009 and March 18, 2010. A schematic representation of the location of the sampling sites within the cave system and the timing of floods (as determined by continuously logging of water levels by on site pressure gauges, Ingenieurbüro Ziegler GmbH), is presented in suppl. Fig. S1. Material for chemical water analysis, and the determination of bacterial abundance and community structure was obtained at every sampling time point (for the results of these analyses see Shabarova et al 2013). Eleven additional water samples (1 l each, during samplings 1,2,3,5 and 7 in pools LC and NGIII, and sampling 7 in pool NS) were collected in Teflon bottles to further analyze the composition of the dissolved organic carbon (DOC) by means of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). Moreover, 13 samples of bacterial biomass were filtered onto Sterivex-GP membrane filter units (Merck Millipore), for DNA extraction and subsequent analysis of 16S rRNA genes by means of terminal fragments length polymorphism (T-RFLP), clone libraries (see Shabarova et al 2013) and 454 tag pyrosequencing sequencing (this study). All samples were delivered to the lab within 24 h after sampling. The water samples for FT-ICR-MS were frozen at -20 °C.

DNA extraction

DNA extraction from the biomass samples collected on Sterivex-GP cartridges was performed within 2 days after sampling using the UltraClean® Water DNA isolation kit (MO BIO Laboratories, Inc.). For more details on the extraction procedure see (Shabarova et al 2013). Eight DNA extracts corresponding to sampling 1, 3, 5, and 7 from pool NGIII, samples 1 and 7 from pool LC and sample 7 from pool NS (suppl. Fig. S1) were selected for 454 tag pyrosequencing of 16S rRNA genes. For this purpose, subsamples of 300 µl of DNA suspension in DNA-free water (final concentration, 4-10 ng μl^{-1}) were delivered to and further processed by Research and Testing Laboratory, Inc. (Lubbock, TX, USA).

16S rRNA tag sequence analysis

Partial 16S rRNA gene sequences were obtained from 454 pyrosequencing (Roche FLX platform) using primers 341F (Muyzer et al 1993) and 907R (Lane et al 1985). Raw data (95516 reads, mean raw read length 571 base pairs) were processed by a custom-made pipeline on a local computer cluster consisting of 16 units (each equipped with an 8 core AMD FX-8150 CPU, 16 GB RAM and a 128 GB SSD hard disk) and a separate control workstation. The program was realized in DELPHI and runs under Windows 7.

Reads were denoised at the level of flowgrams (Quince et al 2011) by first truncating flowgrams at the second instance of a flow signal intensity ranging between 0.5 and 0.7, indicative of a low quality position. This strategy was chosen to avoid excessive data reduction, since a considerable number of flowgrams had a first case of such an intensity range at their very beginning but was of high quality thereafter. In addition, flowgrams were truncated if there were 4 consecutive flows with values <0.5, i.e., if there was no clear base calling at a particular position. Finally, if flowgrams were <360 after these truncation steps (smaller than half the average flow length), they were also discarded. A distance matrix was calculated from the remaining high-quality flowgrams. The flowgrams were preclustered by a complete linkage approach at a similarity level of 97%, corresponding to a similarity of approximately 98.5% at the level of sequences. The central flowgrams from each cluster, defined by a minimal average distance to all members of the cluster, were subsequently chosen and translated into DNA. Chimera checking was performed at the level of DNA sequences using UCHIME (Edgar et al 2011). Altogether our various quality filtering strategies reduced the number of raw reads by 28.5%. The preclustering of flowgrams further reduced this number to 9768 sequences.

After the pairwise alignment (Needleman-Wunsch algorithm) of the representative DNA sequences from flowgram preclustering, a distance matrix was calculated and operational taxonomic units (OTUs) were produced by complete linkage at similarity levels of 97%, 95%, 90%, 85%, 80% and 75%. For the taxonomic assignment of 97% OTUs we developed the following strategy: First, the central sequence from each OTU was selected and the most closely related sequence in the SILVA reference data base release 109 (Pruesse et al 2007) was identified by means of pairwise alignment. Next, the taxonomic information of this reference sequence was assigned to the respective OTU, in addition to the information about its similarity to the query sequence. OTUs were allocated to larger taxonomic entities based on this similarity value (e.g., to the level of genera only if their similarity to the reference sequence was >95%).

Diversity analysis

For the estimates of diversity and community structure, the number of sequences in each sample was normalized, i.e. samples were rarefied to 6297 randomly chosen sequences. OTU richness and the diversity estimators CHAO, and ACE were calculated from these normalized data by a custom made software at different cutoffs of OTU identity (97% - 85%).

To assess the importance of environmental filtering vs. mass effects in the studied habitats we tested for disproportionally high coexistence of closely related taxa within each sample against the null hypothesis of randomized interspecies distances. Information about the depth of phylogenetic branching was generated by calculating both, the Mean Nearest Taxon Distance (MNTD) and the Mean Pairwise Distance (MPD) (Webb et al 2008) from distance matrices of normalized samples using the Picante R package (Kembel et al 2010). The shuffling of distance matrix labels across all taxa was used as null model (Picante function "taxa.labels", 9999 randomizations). Significance was estimated by means of the (abundance weighted) Nearest Taxon Index (NTI) for MNTD and the Net Relatedness Index (NRI) for MPD (Webb et al 2002). These parameters correspond to the number of standard deviations (SD) between MPD or MNTD and the means of the distributions of randomized distances. Comparisons of the phylogenetic relatedness between pairs of assemblages were assessed by the Beta Nearest Taxon Index (βNTI) and the Beta Net Relatedness Index (βNRI), using Picante in combination with a custom R script (Stegen et al 2012, Webb et al 2011).

In addition, a parameter termed Mean Community Distance (MCD) was developed in order to estimate the overall relatedness (distance) of sequences within a sample, in analogy with calculating the total branch lengths of a phylogenetic tree. MCD of each normalized sample was calculated by first performing a recursive clustering (Complete Linkage algorithm) at ascending distances until all sequences were included in a single cluster, by summing up the distances between newly formed clusters at each step of the recursion, and by dividing the final sum by the total number of analysed sequences. In addition, the numbers of sequences in distance ranges of 0-3%, 3-6%, 6-10% and >10% was calculated. Total MCD values and the MCD distributions were used to compare the phylogenetic changes of communities in the initial and final samples from different floods. In order to assess the significance of differences in MCD between samples with different water residence time, 1000 sets of 6297 sequences were randomly selected from the pool of all normalized samples for MCD calculations. Differences between samples were deemed

significant if they were >2 SD of the distribution of MCD values from the randomized assemblages.

FT-ICR-MS analysis of DOC composition

For analysis on the FT-ICR-MS, all extracts were diluted to a DOC concentration of 20 ppm in a 1:1 mixture of ultrapure water and methanol. Analysis was carried out on a 15 Tesla Solarix FT-ICR-MS (Bruker Daltonics, Bremen, Germany) applying electrospray ionization (ESI) in negative mode (500 accumulated scans, 2 s ion accumulation time, mass window: 153 to 2000 Dalton). The spectra were calibrated based on an internal calibration list using the Bruker Daltonics Data Analysis software package. Molecular formulae were assigned to peaks of exported mass lists using in-house Matlab routines. Formula assignment was based on the following criteria: $C_{1-40}H_{1-n}O_{1-n}N_{0-2}S_{0-1}$, minimum H/C 0.3, maximum O/C 1, mass tolerance 500 ppb. Data interpretation focused on the mass range 200-800 Dalton, accounting for >95% of total peak area. Masses only detected in single samples with peak areas <10 times signal to noise were excluded. Detected masses associated with blanks were removed (212 of 5560 peaks), and the remaining peak areas were normalized to total peak area of each sample.

Molecular formulae were split into 6 categories according to their Aromaticity Index (AI_{mod} (Koch and Dittmar 2006)) and other chemical properties as follows: combustion-derived polycyclic aromates (PCAs, $AI_{mod} > 0.66$), soil-derived polyphenols and PCAs with aliphatic chains ($0.66 >= AI_{mod} > 0.50$), soil-derived "humics", i.e. phenolic and highly unsaturated compounds ($AI_{mod} <= 0.50$ and H/C < 1.5), unsaturated aliphatic compounds (2.0 > H/C >= 1.5), saturated fatty and sulfonic acids, carbohydrates including amino sugars (H/C >= 2.0 or O/C >= 0.9), and peptides (as previous group, but containing N).

The differences between the initial and final samples from all three flood events were calculated for the sums of peaks within specific classes of molecular formulae. These classes were defined by the number of C or O atoms of individual formulae, by their O/C ratios, and by their double bond equivalents minus O (DBE-O = 1 + 0.5·× (2C - H + N) - O) (Gonsior et al 2009):

Statistical analysis

The abundances of OTUs in each sample were normalized and agglomerative hierarchical clustering was then conducted using Pearson correlation (complete linkage algorithm) with the software XLSTAT 11 (Addinsoft) or with SPSS Statistic 20 (IBM). Clustering of samples according to relative abundances of DOC compounds was performed in the same manner. In addition, the number of shared OTUs between pairs of samples and θ values (Yue and Clayton 2005) were calculated. In order to assess the robustness of clustering results the matrices of θ values and from agglomerative hierarchical clustering were compared by Mantel tests.

To identify the most important variables (OTUs, molecular formulae) responsible for the observed clustering patterns we utilized the R package Boruta, which is a feature selection algorithm that, in turn, implements a 'random forest' algorithm (Kursa and Rudnicki 2010). A default confidence level of 0.999 was set. In order to reduce noise, small OTUs and rare molecular formulae were excluded prior to the analysis, to the extent that the overall clustering patterns remained unaffected (Aagaard et al 2012).

Results

Extend of novel diversity

In order to assess if the various microbial genotypes in karst cave pools were closely affiliated to already known taxa or if they represented genuinely novel groups we analyzed the distance distribution between the central sequences of each OTU (97% identity) and their closest relative in the SILVA reference database. Almost half of all sequences from the pooled dataset (32027 of 68264) were in OTUs that had <1 % difference to their closest relatives in the SILVA database. However, these OTUs represented <10% of all OTUs (337 out of 4804). The modal class in the OTU distance distribution was at 3-4 % distance, corresponding to 648 OTUs (with 7978 sequences).

The peaks in the similarity distribution of all sequences corresponded to specific maxima of different bacterial groups, and there were clear differences in the distance of individual phylogenetic groups to the known diversity (suppl. Fig. S1). For example, 63% and 62% of sequences affiliated with β-*Proteobacteria* and *Actinobacteria* (26419 and 2292 sequences, respectively) were from OTUs with <1% distance to their closest SILVA relatives. OTUs harboring 42% of all *Bacteroidetes* sequences had distances of 1-2%, while the distribution maximum of sequences from γ-*Proteobacteria* was in OTUs with 3-4 %

distance. Larger distances to known genotypes were found for δ-*Proteobacteria,* as well as for some groups of cultivable (such as *Firmicutes* and *Cyanobacteria*) and non-cultivable bacteria (candidate divisions TM6 and SM2F11). At the extreme, most members from the candidate division OD1 had distances >15% to their closest relatives in the SILVA database. Consequently, the dominance of different phylogenetic lineages changed with distance to known diversity (Fig. 1A, insert): While β-*Proteobacteria* were 38-83 % of all sequences in OTUs with distances <10%, their proportions dropped to 10-12 % at larger distances. On the other hand, candidate division OD1 was the largest phylogenetic group in OTUs with distances >15%, but had no OTUs with distances <1-2%.

Except for *Bacteroidetes* and *Actinobacteria*, there were considerable differences in the proportional taxonomic affiliations of sequences and OTUs in the various phylogenetic lineages (Fig. 1A, inset). Several groups that formed major fractions of all OTUs (α-, γ-, and δ-*Proteobacteria*) were clearly underrepresented in terms of sequences, and this was also the case for more rare bacterial divisions such as *Acidobacteria*, *Elusimicrobia* and candidate division SM2F11 (each >2% of all OTUs). By contrast >50% of sequences belonged to the class β-*Proteobacteria,* but these bacteria were considerably less dominant in terms of their proportion of total OTUs.

Core community of the studied cave pools

Only 74 OTUs (1.5% of all OTUs) were shared between all 8 samples, including the 20 largest ones (Fig.1B). Consequently, almost two thirds (64%) of all sequences were affiliated with these 'core OTUs'. In contrast to the whole community there was a very similar proportional distribution of core OTUs and sequences across different phylogenetic groups (Fig.1B, inset). Moreover, the majority of core OTUs and sequences therein were very closely related to sequences in the SILVA database (Fig.1C). The largest group of core OTUs were β-*Proteobacteria* (56 OTUs harboring 34047 sequences) followed by *Bacteroidetes* (8 OTUs with 6544 sequences) and *Actinobacteria* (5 OTUs with 2203 sequences). The most closely related sequences of most core OTUs in the SILVA database originated from various aquatic environments (Tab.1). In particular, the closest relatives of 35% of the core OTUs (representing >50% of the core sequences) have been found in karst groundwater.

In order to better assess the phylogenetic depth of the core community (i.e., the numbers and proportions of core 'species', 'genera', 'orders', 'classes', and 'phyla') we calculated OTUs at different levels of sequence identity. The numbers of larger taxonomic

entities specific to the core assemblage showed a distinct minimum at the approximate level of 'genera' (95% identity), and remained almost unchanged until the approximate level of phyla (75% identity) (Fig 2). Moreover, while the fraction of core OTUs within all OTUs increased with declining OTU identity, still only 40 of 116 of phylum-like OTUs were shared between all samples (Fig.2).

Changes of microbial diversity and community structure with water residence time

In addition to changes in water chemistry we also observed a reduction of microbial diversity after prolonged water residence times in initial vs. final samples from all three flooding events. This was apparent for OTU richness, but also for two estimators of total diversity ACE and CHAO 1, and at OTU identity definitions of 97% and 90% (Fig. 3 A, B).

NTI indices showed significantly smaller phylogenetic distances among co-occurring OTUs (97% identity) in all samples compared to those estimated by a randomization approach (range: 5.7 to 9.1). NRI values also indicated a significant accumulation of closely related taxa (range: 6.2 to 10.6). All pairwise comparisons of communities yielded highly significant values of βNTI (range, 7.0 to 10.6), whereas βNRI was not significant in the majority of cases (range, 0 to 4). In particular, the microbial assemblage in NS 1 did not significantly relate to any other sample (βNRI values <1.6). MCD values of all samples except NS 1 were significantly lower (range: 5.4 to 7.7) than of 1000 randomly drawn subsamples ($MCD_{randomized}$ = 8.3, SD = 0.14), whereas MCD of NS 1 (9.3) was significantly higher than the average of distances obtained by chance. The microbial assemblages moreover became significantly more similar (i.e., differences >2 SD of the randomized samples) with prolonged water residence time (MCDs of LC1 vs. LC7: 6.2 to 5.5; NGIII1 vs. NGIII3: 7.8 to 5.5; NGIII5 vs. NGIII7: 6.5 to 5.7). The distributions of sequences over different distance classes revealed an increase in the numbers of sequences that clustered at an identity level of 97% and a simultaneous reduction of sequences with higher phylogenetic distances (Fig.3C).

Separation of microbial communities according to water residence time and flood events

Agglomerative hierarchical clustering based on the relative abundances of sequences within OTUs (97% identity) displayed a similarity of samples with comparable water residence times (Fig.4A): Three clusters with short, intermediate, and long residence times could be separated. An identical clustering pattern was obtained from the pairwise

comparison of community similarity using θ values (Mantel test: r=0.94; P< 0.0001). Random forest analysis could identify a set of 7 OTUs that were most strongly responsible for this pattern (suppl. Tab. 1). Of these OTUs, 5 were affiliated with β-*Proteobacteria.* Accordingly, a similar clustering pattern with high correlation to the clustering of all sequences (Mantel test: r=0.94; P<0.0001) was obtained if only β-*Proteobacteria* were considered (Fig. 4B). By contrast, a separation into samples with short and long residence times only was observed if sequences affiliated with β-*Proteobacteria* were excluded (Fig. 4C). This dichotomy of early and late samples was moreover significantly correlated with a clustering of OTUs affiliated with the γ-*Proteobacteria* only (data not shown, Mantel test: r=0.65; P<0.0001) or of the most abundant OTUs of *Bacteroidetes* (data not shown, Mantel test: r=0.92; P<0.0001). By contrast, a clustering of samples using all OTUs associated with *Bacteroidetes* resulted in a clearly different pattern, i.e., samples from the first flood event were significantly different from those of the following two floods (Fig.5A). Random forest analysis suggested that key *Bacteroidetes* defining the first flooding event were related to *Cytophagaceae* (1 OTU) and *Cryomorphaceae* (4 OTUs), whereas 5 OTUs affiliated with *Flavobacteriaceae* and 1 OTU of unclassified subsurface *Sphingobacteriales* were characteristic of the second and third floodings (suppl. Tab. 2),

Analysis of DOC composition

Samples were clustered into three groups (Fig.5B) that matched with the different flood events (first flood: pool LC, 18.11.2009; second flood: pool NGIII, 14.12.2009; third flood: pool NGIII, 31.12.2009, suppl. Fig S1). Despite a partial flood event on 3.04.2010 (suppl. Fig S1) sample NS7 showed high similarity with sample NGIII7. Random forest analysis was conducted to obtain the DOC compounds most relevant for discriminating the observed clustering. It yielded 33 ‘fingerprint’ molecular formulae only composed of H, C and O that showed little variance within samples of each flooding event. The two subsets of DOC compounds discriminating between floods had maximal peak areas in the first and third flooding event, respectively, and showed intermediate values in the second one. They could be clearly distinguished according to their O/C ratios and molecular masses (Fig. 5C).

We moreover observed clear changes of DOC quality that were similar for all three floods in samples with prolonged water residence time. The differences of molecular formulae between early and late samples (separately calculated for classes defined by numbers or ratio of atoms) showed similar trends for the number of C and O atoms: An enrichment of shorter molecules together with a reduction of the abundances of longer molecules was observed (Fig. 6). The reduction of molecule length was accompanied by a

343 drop of oxidation levels especially in sample pairs separated by longer water residence times
344 (NGIII5-NGIII7 and LC1-LC7). In addition the composition of molecular formulae between
345 samples also changed according to their DBE-O values: except for molecular formulae with
346 DBE-O >12, compounds with higher DBE-O were more common in the final samples.

347 The highest proportion of DOC in all 11 samples as discriminated by FT-ICR-MS
348 analysis was attributed to category III (soil-derived 'humic' compounds, 70.2 to 76.1% of total
349 peak area). Category II (soil-derived polyphenols and PCAs with aliphatic chains)
350 contributed 14.1 to 16.8% to total peak area. DOC from category I (combustion-derived
351 PCAs) and IV (unsaturated aliphatic compounds) represented 6.3 to 9.5% and 2.7 to 3.8% of
352 total peak area, respectively. The summed intensities of DOC components from categories V
353 and VI (saturated fatty and sulfonic acids, carbohydrates and peptides) did not exceed 0.5%
354 of total peak area in any sample. A comparison between initial and final samples showed an
355 increase in the relative peak intensities of molecular formulae from categories I and II after
356 all three flooding events (by 9 ± 4.6% and 3.5 ± 2.6%, respectively). By contrast, peak area
357 associated with categories III and IV decreased during residence time.

Discussion

'Novel' vs. 'known' microbial diversity in karst ponds

Exponential growth of public databases of 16S rRNA gene sequences has led to a more complete coverage of the 'known' bacterial diversity (Nakamura et al 2012), as also reflected in the discovery of numerous deeply branching bacterial lineages without cultivated representatives, e.g. Candidate Division OD1 (Harris et al 2004). It is thus conceivable that only little novel diversity might be revealed in some environmental samples by NGS and that instead there is increasing resampling of 'known' diversity. We used the most precise available approach (pairwise alignment) to determined the distances of the central sequences within each OTU to their closest known relative in a well-curated database (Pruesse et al 2007) (Fig. 1A) in order to select a meaningful precision of taxonomic assignment, i.e., if OTUs should be classified at the level of genera, families, classes, etc. This provided a more information-rich alternative to a simple categorization of 'exotic' OTUs (i.e., with large distances to their closest reatives) as 'unclassified' (Hong et al 2011, Pronk et al 2009) in that it revealed differences between and within the individual phylogenetic lineages in karst water: For example, substantially more OTUs of δ-*Proteobacteria* had larger distances to known genotypes than of β-*Proteobacteria* (Supp. Fig 2). The two largest δ-proteobacterial OTUs (approximately one third of all sequences from this lineage) were closely related to a known taxon (*Bdellovibrio* sp., distances <1%), whereas each of the numerous small OTUs affiliated with *Myxococcales* represented a novel genus (average distances of 7%). Genuinely novel microbial diversity was disproportionally high in genotypes related to Candidate Division OD1, which were the majority of all sequences and OTUs with distances >15% to 'known' SILVA genotypes (Fig. 1A). Two OTUs of this lineage had closest relatives from a similar subsurface karst environment (Pronk et al 2009), albeit with distances of 15% and 18%, respectively. Thus, while representatives of OD1 have been obtained from a variety of habitat (Borrel et al 2010, Briée et al 2007, Chouari et al 2005, Elshahed et al 2005), it appears that the 'rare biosphere' of endokarst environments harbors exotic and potentially endemic lineages related to this group. Interestingly, the OD1 related sequences in our samples also were only distantly related to each other, and only approximately half of them could be clustered into OTUs even at an identity cutoff of 85%. Unfortunately, the considerable phylogenetic distances to known representatives and to each other render it difficult to deduce physiological properties of these bacteria from the metagenomic information available for other representatives of OD1 (Wrighton et al 2012).

The 'core' microbial assemblage in the studied karst water system

While well-delineated small or singleton OTUs may be useful for revealing the phylogenetic distribution of 'novel' diversity, it is equally important to identify abundant microbial genotypes that are likely most relevant for system processes (Pedrós-Alió 2006). The concept of the 'autochthonous microbial endokarst community' (AMEC) has been proposed to accommodate the conspicuous invariance of bacterial community composition in karst spring water irrespective of season and hydrological events (Farnleitner et al 2005). Our definition of a 'core' assemblage overlaps with the idea of the AMEC, yet it narrows the focus to those bacteria that can persist or thrive in temporary oxygenated karst water habitats. By restricting the core to the subset of large OTUs (Fig. 1B) that were present in all samples, we conservatively excluded members of the AMEC that might have originated from other endokarst environments with, e.g., anoxic or sulfuric conditions (Engel 2010b), together with other, non-persistent bacteria introduced from interconnected habitats. This interpretation is supported by the observation that only a minority of all bacterial 'phyla' found in karst water were shared between all samples (Fig. 2).

A large set of core OTUs (a clade of *Comamonadaceae* representing >50% of core sequences) were most closely related (0.3-8.5%) to two bacterial 16S rRNA gene sequences from another study of karst water bacteria in the Swiss Alps (Pronk et al 2009). This may be regarded as a first indication for the existence of regional occurrence patterns of particular bacterial genotypes in this habitat (Fierer 2008, Hanson et al 2012). In addition, the considerable microdiversification of bacteria from this lineage (Shabarova et al 2013) suggests adaptive radiation within the endokarst environment. The majority of the remaining core OTUs were affiliated with genotypes from various aquatic habitats (including snow), but also from soil (Fig. 1C, Table 1), possiby reflecting the high connectedness of karst systems at a landscape level, and the ability of some more euryoecious bacterial taxa to survive in the subsurface environment. An intriguing example for typical inhabitants of surface freshwaters that were present in all our samples are a set of OTUs that were virtually identical (<0.4% distance) to known genotypes from the AcI-A tribe of *Actinobacteria* (Eckert et al 2012, Newton et al 2011, Salcher et al 2010). While total bacterial numbers tended to decline (Shabarova et al 2013), the proportion of sequence tags from this lineage increased by two- to four-fold after flooding events (data not shown). Recently, AcI *Actinobacteria* bacteria have been shown to utilize DOC species that are produced during the degradation of chitin and of bacterial cell walls (Beier and Bertilsson 2011, Cloud-Hansen et al 2006, Eckert et al 2013), which may help to explain why they appeared to proliferate under the heterotrophic conditions of the epiphreatic pools.

Selective processes acting on the studied microbial communities

OTU richness and other estimators showed a decline of diversity in all pools with increasing time after flood events and at several levels of phylogenetic resolution (Fig. 3 A, B). This supports our earlier hypothesis derived from the analysis of two 16S rRNA gene clone libraries from pool NGIII only (Shabarova et al 2013). Moreover, Fig. 3C reflects the specific reduction of singletons over time, i.e. of sequences that could not be clustered at an identity level of 97%. Singletons are often excluded for analytical or conceptional reasons (Kunin et al 2010), but nevertheless represent a valid cross-section through the 'rare biosphere' (Huse et al 2010). While there may be stochastic fluctuation in the proportions of singleton OTUs between samples, their consistent loss speaks for the acting of directed processes on microbial assemblages during their residence in the karst pools. In addition, habitat filtering was also suggested by the significant decline of a parameter for overall phylogenetic community distance (MCD) between early and late post-flooding samples.

NTI and NRI (or βNTI and βNRI, respectively) detect phylogenetic signals of deterministic processes within (or between) microbial assemblages (Barberan and Casamayor 2010, Horner-Devine and Bohannan 2006, Pontarp et al 2012, Stegen et al 2012). In a previous study, subsurface microbial communities were most strongly shaped by habitat filtering in temporally and spatially highly variable environments (Stegen et al 2012). NTI and NRI values were highly significant in all our samples, which is in line with results from a global survey of freshwater habitats (Barberan and Casamayor 2010). Highly significant βNTI values between all sample pairs moreover illustrated that the epiphreatic karst water system as a whole strongly determined the 'tips' of the phylogenetic distributions, i.e., was selective for a set of closely related taxa. This was substantially less apparent at the level of deeper phylogenetic branchings, as reflected by a majority of non-significant values of βNRI.

There was apparent contradiction between the various indicators for the communities obtained during the flooding itself (NS1): NTI and βNTI suggested that the community in the influx was not a random collection of genotypes, but appeared to be already preshaped by deterministic processes, i.e., a disproportional ocurrence of closely related genotypes that were also present in all other samples. By contrast, MCD pointed to a significantly higher proportion of more deeply branching diversity (including singletons) in NS1, which was also supported by non-significant βNRI values between NS1 and all other samples (Kraft et al 2007). We previously suggested that the decrease of microbial diversity during residence in the pools (Fig. 3A, B) would indicate a predominance of mass effects (Crump et al 2007) shaping the original flood water communities (Shabarova et al 2013). This conclusion might

be overly simplistic in view of the current, more detailed analysis. Instead, it is conceivable that the flooding events either collected subsets of OTUs from a variety of habitats with different selective conditions (Engel 2010a), or that water transport through subsurface crevices in itself acted as an environmental filter (Shabarova et al 2013). For example, presumeably methylo- or methanotrophic bacteria related to *Crenothrix* that were prominent in the influx probably originated from an anoxic aquatic niche (Stoecker et al 2006), whereas bacteria closely affiliated with soil-dwelling *Pseudospirillum* spp. were likely introduced from a habitat with different selective properties. The contrasting fate of these two gammaproteobacterial taxa during their residence in the karst pools (decline vs. increase, data not shown) again illustrates the subsequent acting of selective processes.

Transformations of microbial assemblages during residence in karst pools

The similarity analysis of karst water communities by NGS after exclusion of β-*Proteobacteria* (Fig 4C) closely corresponded to previous results obtained by molecular fingerprinting (Shabarova et al 2013). In addition, more detailed results could be obtained for particular phylogenetic groups due to the higher discriminative power of NGS. Most importantly, there was a significant separation of the β-proteobacterial sub-community into three sets of samples with different water residence time after flooding (Fig 4B). This pattern was in parts due to OTUs that initially declined in relative abundances in the 'intermediate' samples and increased thereafter *(e.g.,* OTUs 4662, 4755, 3740; suppl. Tab.1). Interestingly, the development of the total cell numbers of β-*Proteobacteria*, as determined by microscopy, showed an exactly opposite trend (Shabarova et al 2013), i.e. a strong increase of cell numbers in intermediate samples. While the proportions of sequence types after PCR amplification likely do not accurately reflect their relative abundances (von Wintzingerode et al 1997), the combined information from microscopic cell counts and OTU sizes nevertheless suggests that many β-*Proteobacteria* were not or only slightly growing in the pools (as previously shown for *Methylophilaceae,* Shabarova et al 2013), whereas the transient rise of these bacteria was mainly due to a comparatively small set of large OTUs related to *Comamonadaceae.* For example, OTU 2600, which was most closely related to a genotype previously found in karst water (AM991264), formed 12% of all sequence tags in early post-flooding samples and 20% in intermediate ones.

The clustering of *Bacteroidetes* was similar to that of the DOC components in that it separated samples according to flooding events rather than by residence time (Fig. 5A, B). The two largest flavobacterial OTUs co-defining this pattern (Suppl. Table 2; IDs: 1748 and 4136) are related to genotypes from glacial ice and Antarctic environments (Segawa et al

2010, Skidmore et al 2005, Van Trappen et al 2003), which would agree with the establishment of a surface winter cover in the catchment area between the first and second flooding.

Changes of DOC composition between floods and during residence in karst pools

A clear separation of samples according to flooding events (Fig. 5B) pointed to changes in DOC quality of the source water. The composition of DOC in soils and in stream water can, e.g., be affected by frost and snow melt (Agren et al 2010, Haei et al 2012), and it is conceivable that these processes also determined the similarity patterns in the studied karst pools (Fig. 5B). Thirty of the 33 formulae that were most characteristic for this separation (Fig. 5C) were classified as representatives of 'soil derived humic compounds', which formed the largest class of detected DOC components in our samples. High proportions of humic or fulvic substances in karst water have been reported previously (Simon et al 2010).

Dissolved compounds from terrestrial sources can be an important resource for bacterial activity in aquatic habitats (Berggren et al 2009). DOC concentrations substantially decreased during the residence of water in the pools (by up to 20% Shabarova et al 2013), most likely due to microbial activity. This decrease was accompanied by a transformation of particular DOC characteristics (Fig. 6). Both, the smaller molecular size (number of C atoms) and the higher aromaticity index (Koch and Dittmar 2006) of DOC in final samples agree with current concepts about microbial decomposition of organic matter towards more stable and recalcitrant forms (Amon and Benner 1996).

Acknowledgements:

We would like to thank the cavers of SGH Bern especially Diego Sanz, Daniel Burkhalter, Martin Achtman and Stefan Huber for their help in preparing and conducting the sampling campaigns. T.S was supported by a stipend of the Research Commission of the University of Zurich (grant Nr. 57150401).

526 **Legends to Figures**

527 **Figure 1:** (A) Bar chart: Phylogenetic distances of operational taxonomic units (OTUs, 97%
528 identity) and associated sequences from bacteria in subsurface cave pools to the most
529 closely related genotype in the SILVA reference database (release 109). Pie charts:
530 Phylogenetic composition of OTUs and associated sequences in the whole data set (inset),
531 and of sequences within different distance ranges (as denoted by black lines). (B) Bar chart:
532 Size distributions of all OTUs and of the 'core' OTUs (i.e., present in all samples). Pie charts:
533 Phylogenetic composition of bacteria in the 'core' assemblage. (C) Phylogenetic distances of
534 OTUs and associated sequences in the 'core' assemblage to the most closely related
535 genotype in the SILVA reference database. Missing sections in pie charts denote proportions
536 of sequences/OTUs from rare taxa (<2% of abundances).

537 **Figure 2:** Bars: Numbers of all OTUs and of OTUs in the 'core' assemblage at different OTU
538 identity cutoffs. Symbols: Proportions of 'core' OTUs of all OTUs.

539 **Figure 3:** (A), (B) OTU richness and diversity estimators CHAO1 and ACE in initial and final
540 samples from 3 different flood events at different OTU identity cutoffs (97%, 90%). (C)
541 Distributions of intersequence distances within one sample: Differences between initial and
542 final samples from 3 flood events.

543 **Figure 4:** Similarity analysis of microbial communities in different samples by agglomerative
544 hierarchical clustering of OTUs (97% identity cutoff, Pearson correlation, complete linkage).
545 (A) All *Bacteria* (B) *β-Proteobacteria* (C) All *Bacteria* excluding *β-Proteobacteria.*

546 **Figure 5:** (A) Similarity analysis of *Bacteroidetes*. (B) Similarity analysis of DOC composition
547 as determined by FT-ICR-MS analysis. Clustering in (A) and (B) was performed as outlined
548 in Fig. 4. (C) Physicochemical properties of 33 molecular formulae that were the main
549 determinants for the clustering pattern of DOC composition according to Random Forest
550 Analysis.

551 **Figure 6:** Comparison of selected chemical characteristics of DOC components between
552 initial and final samples from 3 flood events: Differences in number of C atoms, O/C ratios,
553 and Aromaticity Index AI_{mod}.

554

555 **Supplementary material:**

556 **Supplementary Figure 1:** Timing and magnitude of flooding events of the studied
557 epiphreatic pools in the Bärenschacht cave system (modified from Fig.1 in Shabarova et al

2013). The left half of the diagram schematically illustrates the hydrological connection between pools and the required rise in the water table for a flooding event to occur. Black circles, location of temperature/pressure sensors; dashed lines: time points of samplings; squares: samples for FT-ICR-MS analysis; asterisks: samples for 454 analysis. Grey background colors represent water originating from different floods, the hatched area corresponds to a partial flood event occurring on March 2 in NS. Ellipsoid inset: magnified depiction of flooding event that occurred in pool NGIII on December 30 after sampling 3.

Supplementary Figure 2: Phylogenetic distances of OTUs (97% identity) and associated sequences to the most closely related genotype in the SILVA reference database (release 109) within different phylogenetic lineages

568 **References:**

569 Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C *et al* (2012). A metagenomic
570 approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* **7:**
571 13.

572 Agren A, Haei M, Kohler SJ, Bishop K, Laudon H (2010). Regulation of stream water
573 dissolved organic carbon (DOC) concentrations during snowmelt; the role of discharge,
574 winter climate and memory effects. *Biogeosciences* **7:** 2901-2913.

575 Amon R, Benner R (1996). Bacterial utilization of different size classes of dissolved organic
576 matter. *Limnology and Oceanography* **41:** 41–51.

577 Amon RMW, Benner R (1994). Rapid cycling of high-molecular-weight dissolved organic
578 matter in the ocean. *Nature* **369:** 549-552.

579 Barberan A, Casamayor EO (2010). Global phylogenetic community structure and beta-
580 diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology*
581 **59:** 1-10.

582 Beier S, Bertilsson S (2011). Uncoupling of chitinase activity and uptake of hydrolysis
583 products in freshwater bacterioplankton. *Limnology and Oceanography* **56:** 1179-1188.

584 Berggren M, Laudon H, Haei M, Strom L, Jansson M (2009). Efficient aquatic bacterial
585 metabolism of dissolved low-molecular-weight compounds from terrestrial sources. *ISME*
586 *Journal* **4:** 408-416.

587 Bertilsson S, Tranvik LJ (2000). Photochemical transformation of dissolved organic matter in
588 lakes. *Limnology and Oceanography* **45:** 753-762.

589 Borrel G, Lehours A-C, Bardot C, Bailly X, Fonty G (2010). Members of candidate divisions
590 OP11, OD1 and SR1 are widespread along the water column of the meromictic Lake Pavin
591 (France). *Archives of Microbiology* **192:** 559-567.

592 Briée C, Moreira D, López-García P (2007). Archaeal and bacterial community composition
593 of sediment and plankton from a suboxic freshwater pond. *Research in Microbiology* **158:**
594 213-227.

595 Chouari R, Le Paslier D, Dauga C, Daegelen P, Weissenbach J, Sghir A (2005). Novel major
596 bacterial candidate division within a municipal anaerobic sludge digester. *Applied and*
597 *Environmental Microbiology* **71:** 2145-2153.

598 Cloud-Hansen KA, Peterson SB, Stabb EV, Goldman WE, McFall-Ngai MJ, Handelsman J
599 (2006). Breaching the great wall: peptidoglycan and microbial interactions. *Nature Reviews*
600 *Microbiology* **4:** 710-716.

601 Crump RC, Adams HE, Hobbie JE, Kling GW (2007). Biogeography of bacterioplankton in
602 lakes and streams of an Arctic tundra catchment. *Ecology* **88:** 1365-1378.

603 Dittmar T, Hertkorn N, Kattner G, Lara RJ (2006). Mangroves, a major source of dissolved
604 organic carbon to the oceans. *Global Biogeochemical Cycles* **20:** GB1012.

605 Eckert EM, Salcher MM, Posch T, Eugster B, Pernthaler J (2012). Rapid successions affect
606 microbial N-acetyl-glucosamine uptake patterns during a lacustrine spring phytoplankton
607 bloom. *Environmental Microbiology* **14:** 794-806.

Eckert EM, Baumgartner M, Huber IM, Pernthaler J (2013). Grazing resistant freshwater bacteria profit from chitin and cell-wall derived organic carbon. *Environmental Microbiology* **in press**.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics (Oxford, England)* **27:** 2194-2200.

Elshahed MS, Najar FZ, Aycock M, Qu CN, Roe BA, Krumhoz LR (2005). Metagenomic analysis of the microbial community at Zodletone Spring (Oklahoma): Insights into the genome of a member of the novel candidate division OD1. *Applied and Environmental Microbiology* **71:** 7598-7602.

Engel A (2010a). Microbial Diversity of Cave Ecosystems. In: Barton LL, Mandl M, Loy A (eds). *Geomicrobiology: Molecular and Environmental Perspective*. Springer Netherlands. pp 219-238.

Engel AS (2010b). *Microbial Diversity of Cave Ecosystems*. Springer-Verlag Berlin, Heidelberger Platz 3, D-14197 Berlin, Germany.

Farnleitner AH, Wilhartitz I, Ryzinska G, Kirschner AK, Stadler H, Burtscher MM *et al* (2005). Bacterial dynamics in spring water of alpine karst aquifers indicates the presence of stable autochthonous microbial endokarst communities. *Environmental Microbiology* **7:** 1248–1259.

Fazi S, Amalfitano S, Piccini C, Zoppini A, Puddu A, Pernthaler J (2008). Colonization of overlaying water by bacteria from dry river sediments. *Environmental Microbiology* **10:** 2760–2772.

Fierer N (2008). *Microbial biogeography: Patterns in microbial diversity across space and time*.

Ford DC, Williams P (2007). *Karst Hydrogeology and Geomorphology*. Wiley.

Gonsior M, Peake BM, Cooper WT, Podgorski D, D'Andrilli J, Cooper WJ (2009). Photochemically Induced Changes in Dissolved Organic Matter Identified by Ultrahigh Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environmental Science & Technology* **43:** 698-703.

Haei M, Oquist MG, Ilstedt U, Laudon H (2012). The influence of soil frost on the quality of dissolved organic carbon in a boreal forest soil: combining field and laboratory experiments. *Biogeochemistry* **107:** 95-106.

Hama T, Yanagi K (2001). Production and neutral aldose composition of dissolved carbohydrates excreted by natural marine phytoplankton populations. *Limnology and Oceanography* **46:** 1945-1955.

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology* **10:** 497-506.

Harris JK, Kelley ST, Pace NR (2004). New perspective on uncultured bacterial phylogenetic division OP11. *Applied and Environmental Microbiology* **70:** 845-849.

Hong P-Y, Wheeler E, Cann IKO, Mackie RI (2011). Phylogenetic analysis of the fecal microbial community in herbivorous land and marine iguanas of the Galapagos Islands using 16S rRNA-based pyrosequencing. *ISME Journal* **5:** 1461-1470.

Horner-Devine MC, Bohannan BJM (2006). Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* **87:** S100-S108.

Huse SM, Welch DM, Morrison HG, Sogin ML (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology* **12:** 1889-1898.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD *et al* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26:** 1463-1464.

Kirchman DL, Dittel AI, Findlay SEG, Fischer D (2004). Changes in bacterial activity and community structure in response to dissolved organic matter in the Hudson River, New York. *Aquatic Microbial Ecology* **35:** 243-257.

Koch BP, Witt MR, Engbrodt R, Dittmar T, Kattner G (2005). Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochimica Et Cosmochimica Acta* **69:** 3299-3308.

Koch BP, Dittmar T (2006). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Communications in Mass Spectrometry* **20:** 926-932.

Kraft NJB, Cornwell WK, Webb CO, Ackerly DD (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *American Naturalist* **170:** 271-283.

Kujawinski EB, Freitas MA, Zang X, Hatcher PG, Green-Church KB, Jones RB (2002). The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Organic Geochemistry* **33:** 171-180.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology* **12:** 118-123.

Kursa MB, Rudnicki WR (2010). Feature Selection with the Boruta Package. *Journal of Statistical Software* **36:** 1-13.

Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences, USA* **82:** 6955-6959.

Langenheder S, Berga M, Ostman O, Szekely AJ (2012). Temporal variation of beta-diversity and assembly mechanisms in a bacterial metacommunity. *ISME Journal* **6:** 1107-1114.

Muyzer G, de Waal EC, Uitterlinden AG (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59:** 695–700.

Nakamura Y, Cochrane G, Karsch-Mizrachi I (2012). The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research* **24:** 24.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011). A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiol Mol Biol Rev* **75:** 14-49.

Pedrós-Alió C (2006). Marine microbial diversity: can it be determined? *Trends Microbiol* **14:** 257-263.

Pontarp M, Canbäck B, Tunlid A, Lundberg P (2012). Phylogenetic Analysis Suggests That Habitat Filtering Is Structuring Marine Bacterial Communities Across the Globe. *Microbial Ecology* **64:** 8-17.

Pronk M, Goldscheider N, Zopfi J (2009). Microbial communities in karst groundwater and their potential use for biomonitoring. *Hydrogeol J* **17:** 37–48.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J *et al* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35:** 7188–7196.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011). Removing noise from pyrosequenced amplicons. *BMC bioinformatics* **12:** 38.

Salcher MM, Pernthaler J, Posch T (2010). Spatiotemporal distribution and activity patterns of bacteria from three phylogenetic groups in an oligomesotrophic lake. *Limnology and Oceanography* **55:** 846-856.

Segawa T, Takeuchi N, Ushida K, Kanda H, Kohshima S (2010). Altitudinal Changes in a Bacterial Community on Gulkana Glacier in Alaska. *Microbes Environ* **25:** 171-182.

Shabarova T, Widmer F, Pernthaler J (2013). Mass effects meet species sorting: transformations of microbial assemblages in epiphreatic subsurface karst water pools. *Environmental Microbiology* **in revision**.

Simon KS, Pipan T, Ohno T, Culver DC (2010). Spatial and temporal patterns in abundance and character of dissolved organic matter in two karst aquifers. *Fundam Appl Limnol* **177:** 81-92.

Skidmore M, Anderson SP, Sharp M, Foght J, Lanoil BD (2005). Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. *Applied and Environmental Microbiology* **71:** 6986-6997.

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR *et al* (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences, USA* **103:** 12115–12120.

Stegen JC, Lin X, Konopka AE, Fredrickson JK (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME Journal*.

Stoecker K, Bendinger B, Schöning B, Nielsen PH, Nielsen JL, Baranyi C *et al* (2006). Cohn's Crenothrix is a filamentous methane oxidizer with an unusual methane monooxygenase. *Proceedings of the National Academy of Sciences, USA* **103:** 2363-2367.

Van Trappen S, Mergaert J, Swings J (2003). Flavobacterium gelidilacus sp. nov., isolated from microbial mats in Antarctic lakes. *International Journal of Systematic and Evolutionary Microbiology* **53:** 1241-1245.

738 von Wintzingerode F, Gobel UB, Stackebrandt E (1997). Determination of microbial diversity
739 in environmental samples: pitfalls of PCR-based rRNA analysis. *Fems Microbiol Rev* **21:**
740 213-229.

741 Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002). Phylogenies and community
742 ecology. *Annu Rev Ecol Syst* **33:** 475-505.

743 Webb CO, Ackerly DD, Kembel SW (2008). Phylocom: software for the analysis of
744 phylogenetic community structure and trait evolution. *Bioinformatics* **24:** 2098-2100.

745 Webb CO, Ackerly DD, Kembel SW (2011). Phylocom: Software for the analysis of
746 phylogenetic community structure and character evolution (with phylomatic and ecoevolve).
747 *Phylocom Software*.

748 Wrighton KC, Thomas BC, Sharon I, Miller CS, Castelle CJ, VerBerkmoes NC *et al* (2012).
749 Fermentation, Hydrogen, and Sulfur Metabolism in Multiple Uncultivated Bacterial Phyla.
750 *Science* **337:** 1661-1665.

751 Yue JC, Clayton MK (2005). A similarity measure based on species proportions.
752 *Communications in Statistics: Theory and Methods* **34:** 2123–2131.

753 Zhou YX, Kellermann C, Griebler C (2012). Spatio-temporal patterns of microbial
754 communities in a hydrologically dynamic pristine aquifer. *FEMS Microbiology Ecology* **81:**
755 230-242.

Table 1: Origin of genotypes that were most closely related to the core OTUs of karst water samples

Habitat of closest neighbor	**Number of OTUs**	**Number of sequences**
Soil or plant root environment	9	2335
Aquatic environment	**65**	**41404**
Ice, snow, glacier	4	2235
Freshwater planktonic	**39**	**28327**
Karst groundwater	26	23396
Lakes	10	3411
Rivers, streams	3	1520
Freshwater biofilms	3	463
Contaminated fresh water	5	1121
Estuarine	8	5009
Marine	1	161
Geothermal	1	1662
Activated sludge	4	2426
Total	**74**	**43739**

Supplementary Table 1: OTUs that were selected by Random Forest analysis as most significantly responsible for the observed clustering pattern in Fig. 4A, and their mean contributions to all sequences (%) in the three groups of samples.

OTU ID	Acc-Nr	Distance	Phylum	Order	Family	Genus	Early	Intermediate	Late
3038	AB504954	3.5	*Gammaproteobacteria*	*Methylococcales*	*Crenotrichaceae*	*Crenothrix*	0.79	0.70	0.04
2831	FJ712609	1.3	*Betaproteobacteria*	*Hydrogenophilales*	*Hydrogenophilaceae*	*Thiobacillus*	2.18	0.64	0.09
3979	CU921522	0.0	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Acidovorax*	2.00	1.59	0.30
3186	EF590017	5.4	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*		0.00	0.07	1.61
3740	AM991264	3.1	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*		0.42	0.27	0.64
4755	AM991264	4.7	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*		3.45	0.82	2.34
4662	AM991232	0.8	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*		11.28	6.59	16.69

Supplementary Table 2: OTUs that were selected from all OTUs affiliated with *Bacteroidetes* by Random Forest analysis as being most significantly responsible for the observed clustering pattern in Fig. 5A, and their mean contributions to all sequences (%) in the two groups of samples

OTU ID	Acc-Nr	Distance	Order	Family	Genus	Flood 1	Flood 2+3
2047	EU801738	1.5	*Cytophagales*	*Cytophagaceae*	*Arcicella*	0.03	0.00
4244	AF289149	0.4	*Cytophagales*	*Cytophagaceae*	*Arcicella*	0.19	0.05
3608	DQ501343	1.0	*Cytophagales*	*Cytophagaceae*	*Arcicella*	0.02	0.00
2235	AB619698	4.1	*Cytophagales*	*Cytophagaceae*	*Flexibacter*	0.01	0.00
3373	EF471629	6.7	*Flavobacteriales*	*Cryomorphaceae*	*Fluviicola*	0.04	0.00
1029	GU233829	1.9	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	0.02	0.44
1748	AB545735	1.1	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	0.16	2.68
4110	GU233829	1.9	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	0.01	0.10
4136	GU233829	1.9	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	0.22	2.78
4615	AM934634	2.2	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	0.00	0.11
284	AM991241	2.0	*Sphingobacteriales*	PHOS-HE51		0.00	0.05

Figure 1

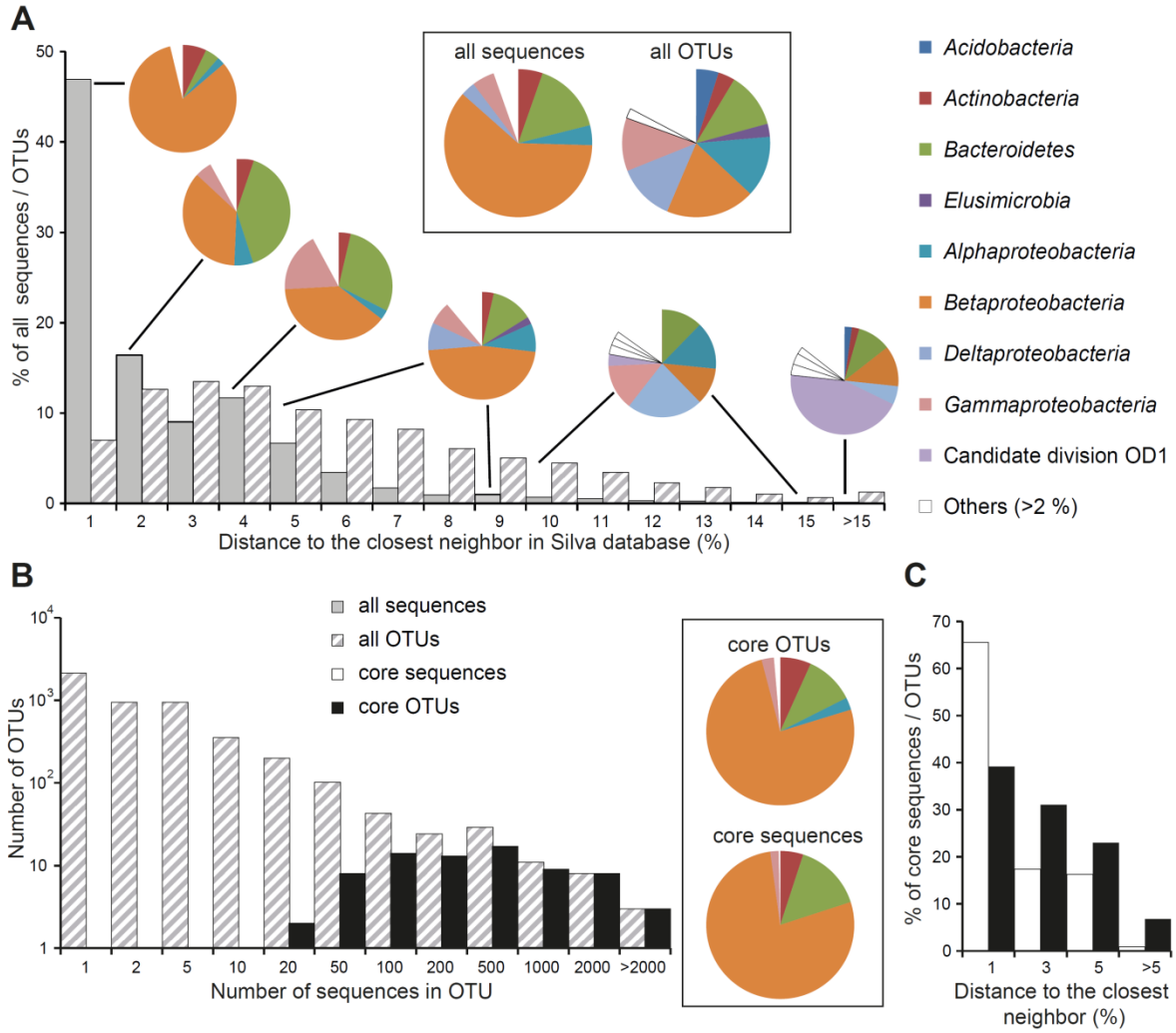

A
% of all sequences / OTUs
Distance to the closest neighbor in Silva database (%)
all sequences
all OTUs
Acidobacteria
Actinobacteria
Bacteroidetes
Elusimicrobia
Alphaproteobacteria
Betaproteobacteria
Deltaproteobacteria
Gammaproteobacteria
Candidate division OD1
Others (>2 %)
B
Number of OTUs
Number of sequences in OTU
all sequences
all OTUs
core sequences
core OTUs
core OTUs
core sequences
C
% of core sequences / OTUs
Distance to the closest neighbor (%)

Figure 2

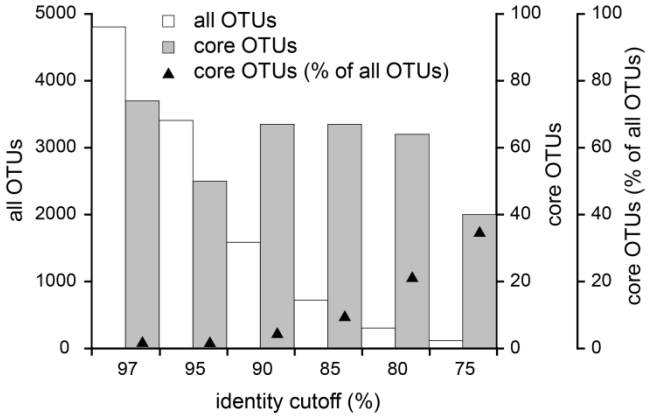

all OTUs
core OTUs
core OTUs (% of all OTUs)
all OTUs
core OTUs
core OTUs (% of all OTUs)
identity cutoff (%)
5000
4000
3000
2000
1000
0
100
80
60
40
20
0
97
95
90
85
80
75

Figure 3

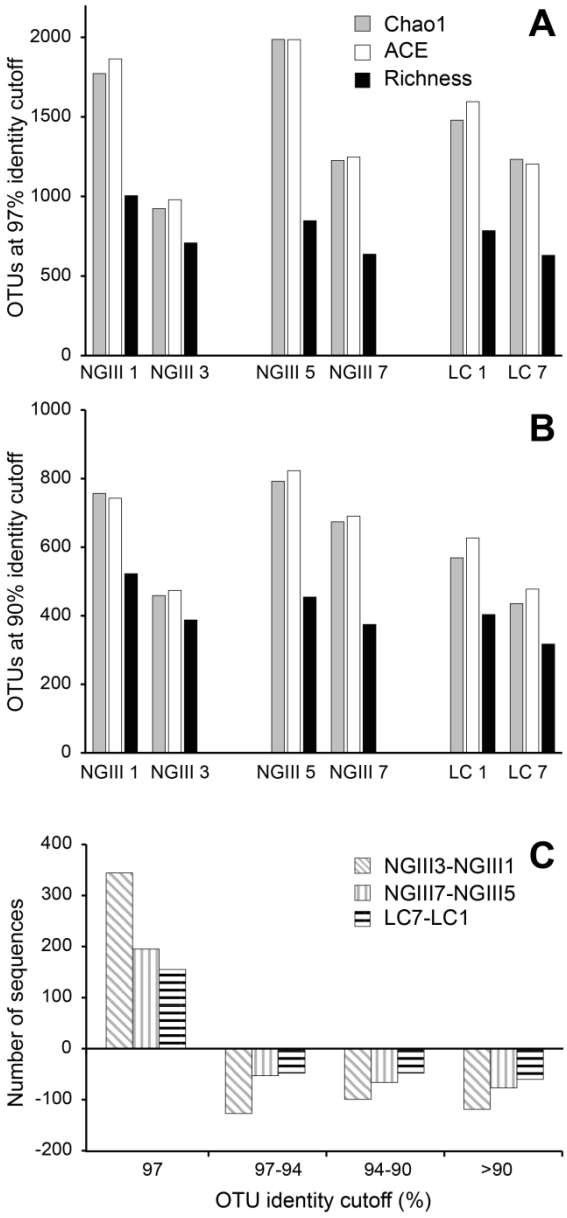

A
Chao1
ACE
Richness
OTUs at 97% identity cutoff
2000
1500
1000
500
0
NGIII 1
NGIII 3
NGIII 5
NGIII 7
LC 1
LC 7
B
OTUs at 90% identity cutoff
1000
800
600
400
200
0
NGIII 1
NGIII 3
NGIII 5
NGIII 7
LC 1
LC 7
C
NGIII3-NGIII1
NGIII7-NGIII5
LC7-LC1
Number of sequences
400
300
200
100
0
-100
-200
97
97-94
94-90
>90
OTU identity cutoff (%)

Figure 4

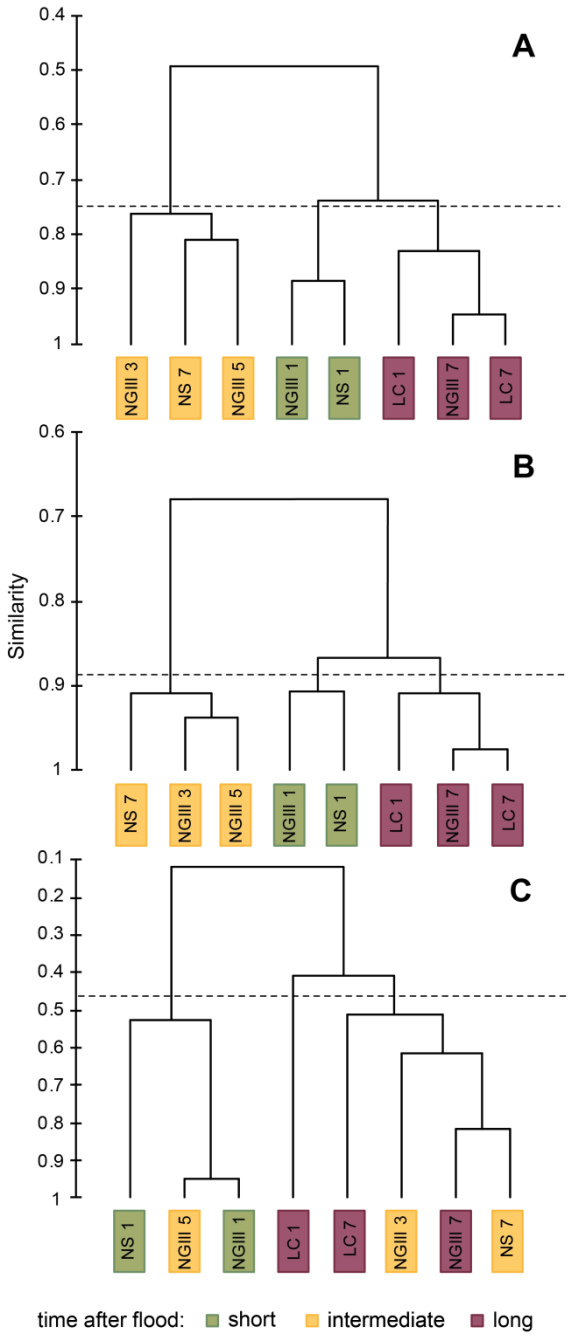

A
B
C
Similarity
NGIII 3
NS 7
NGIII 5
NGIII 1
NS 1
LC 1
NGIII 7
LC 7
time after flood: short intermediate long

Figure 5

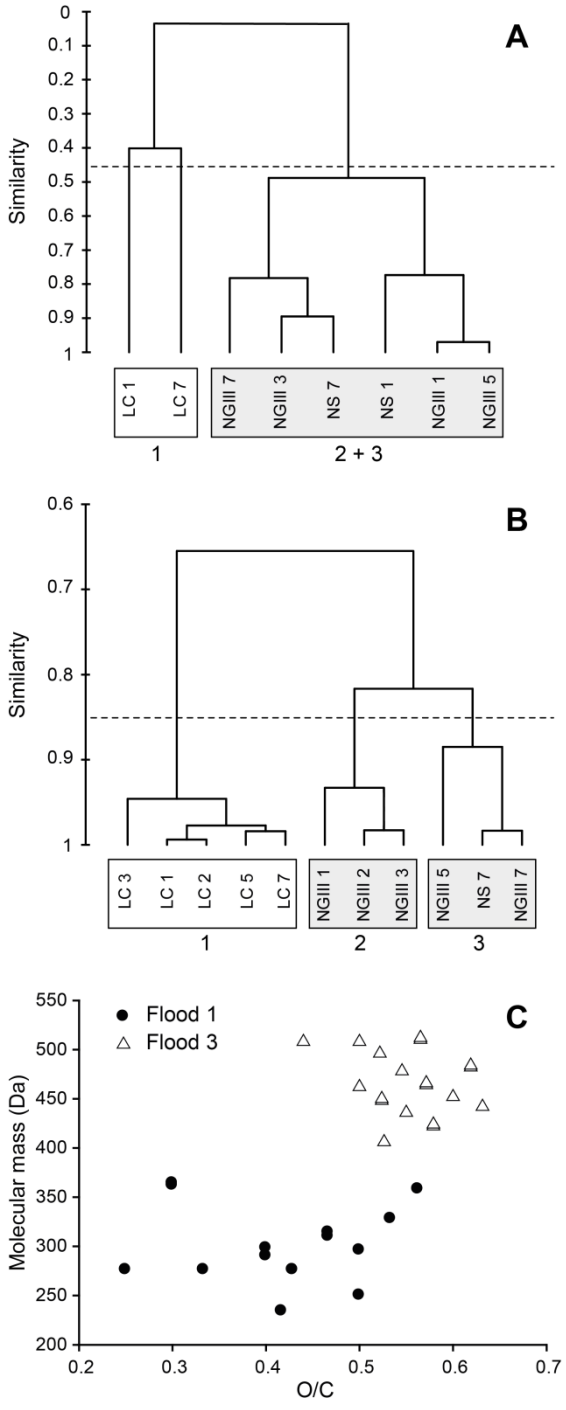
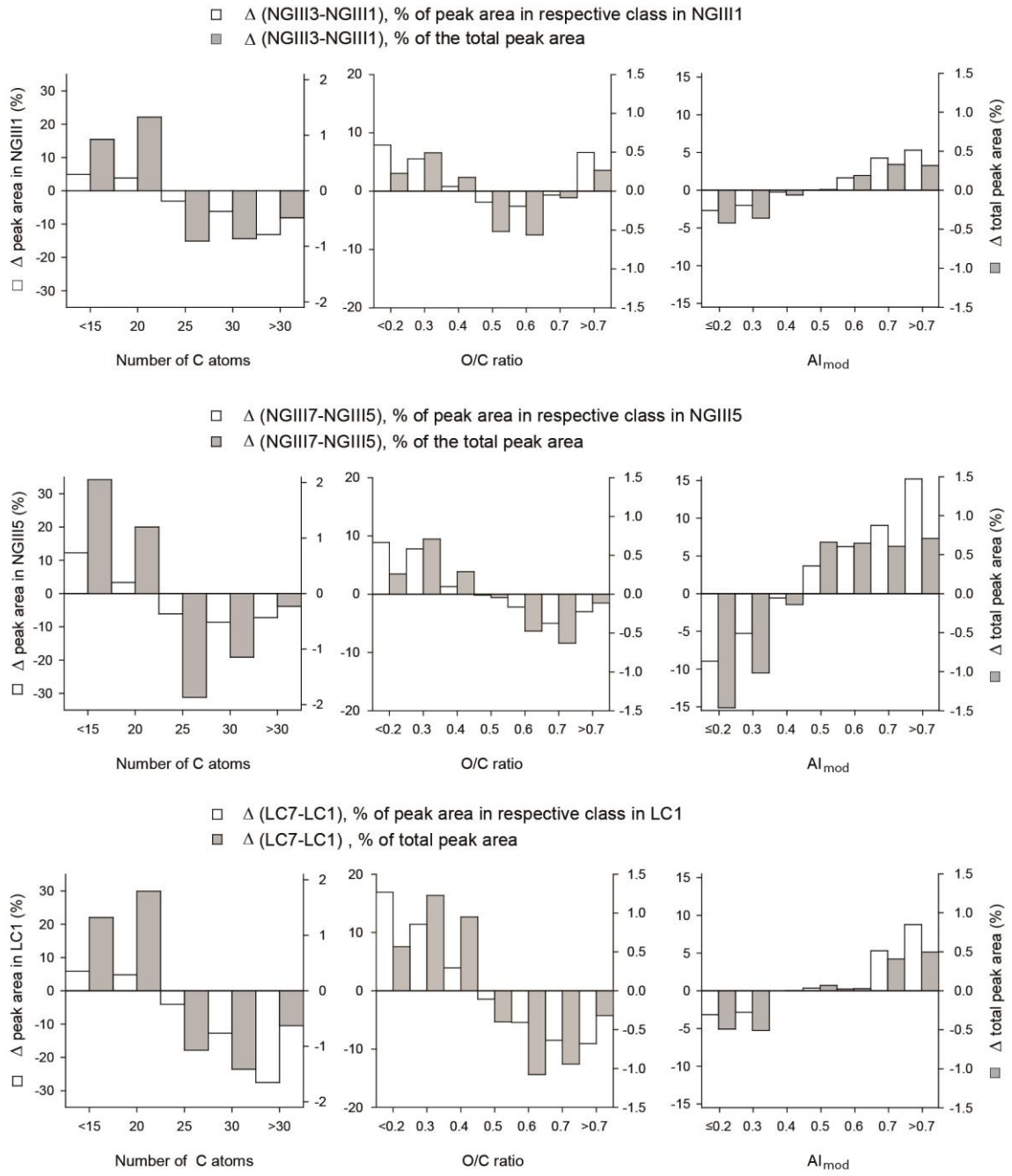
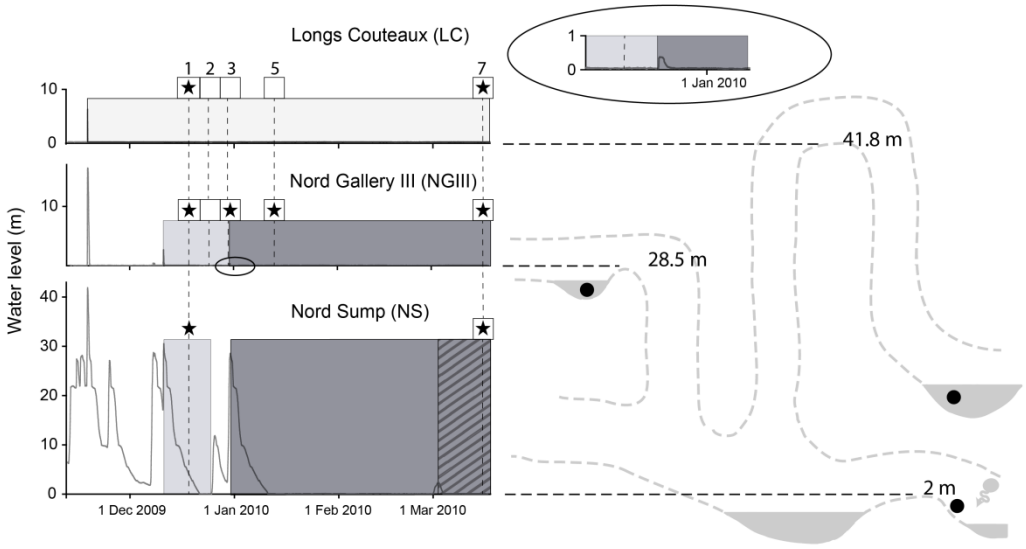

A
Similarity
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
LC 1
LC 7
NGIII 7
NGIII 3
NS 7
NS 1
NGIII 1
NGIII 5
1
2 + 3
B
Similarity
0.6
0.7
0.8
0.9
1
LC 3
LC 1
LC 2
LC 5
LC 7
NGIII 1
NGIII 2
NGIII 3
NGIII 5
NS 7
NGIII 7
1
2
3
C
Flood 1
Flood 3
Molecular mass (Da)
550
500
450
400
350
300
250
200
0.2
0.3
0.4
0.5
0.6
0.7
O/C

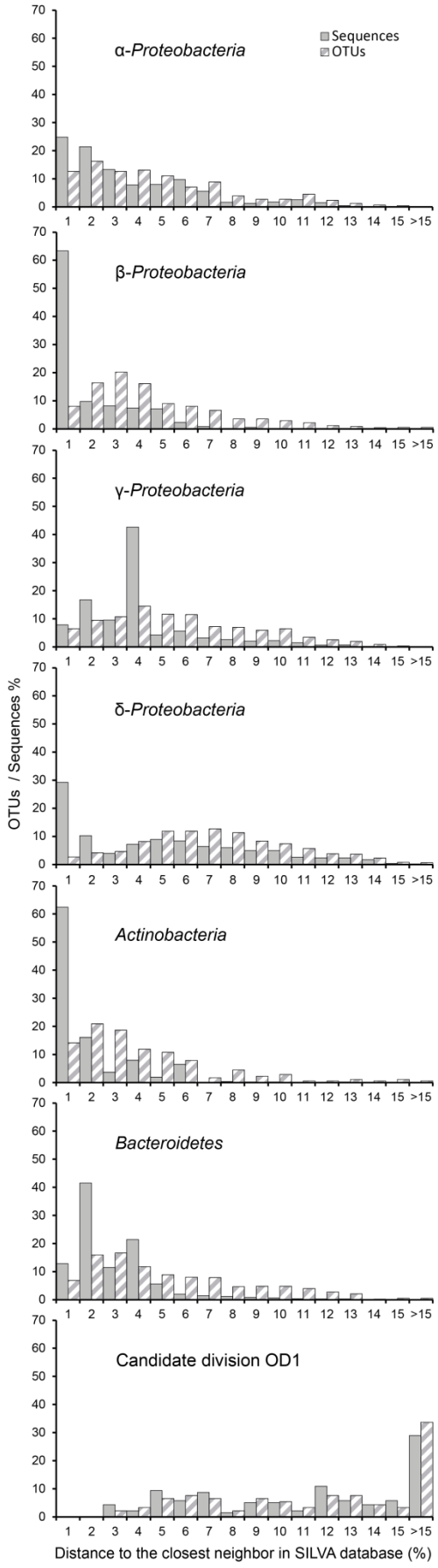
Figure 6


Δ (NGIII3-NGIII1), % of peak area in respective class in NGIII1
Δ (NGIII3-NGIII1), % of the total peak area
Δ peak area in NGIII1 (%)
Δ total peak area (%)
<15 20 25 30 >30
Number of C atoms
<0.2 0.3 0.4 0.5 0.6 0.7 >0.7
O/C ratio
≤0.2 0.3 0.4 0.5 0.6 0.7 >0.7
AImod
Δ (NGIII7-NGIII5), % of peak area in respective class in NGIII5
Δ (NGIII7-NGIII5), % of the total peak area
Δ peak area in NGIII5 (%)
Δ total peak area (%)
<15 20 25 30 >30
Number of C atoms
<0.2 0.3 0.4 0.5 0.6 0.7 >0.7
O/C ratio
≤0.2 0.3 0.4 0.5 0.6 0.7 >0.7
AImod
Δ (LC7-LC1), % of peak area in respective class in LC1
Δ (LC7-LC1) , % of total peak area
Δ peak area in LC1 (%)
Δ total peak area (%)
<15 20 25 30 >30
Number of C atoms
<0.2 0.3 0.4 0.5 0.6 0.7 >0.7
O/C ratio
≤0.2 0.3 0.4 0.5 0.6 0.7 >0.7
AImod

Supplementary figure S1



Supplementary Figure S1



4. General discussion

4.1 Hydrology and water chemistry determining microbiology in karst pools

The exploration of five pools in the Bärenschacht cave revealed a high level of microbial diversity within karst aquatic systems (Manuscript I, Fig. 3 B; Manuscript IV, Fig. 3). Despite the fact that all studied ponds seemed to have comparable richness, clear differences in microbial community characteristics and composition were observed in connection with hydrology.

The vadose pools JI and JII fed by dripping water were characterized by low bacterial abundances ($<10^4$ bacteria ml^{-1}) and distinguishable chemistry. We suggested them to function as collectors of microbial diversity from small individual granular aquifers, which seemed to determine, both, chemistry and microbial community composition in these ponds. Although members of some phyla, in particular of β-*Proteobacteria* and *Actinobacteria,* were abundant in both pools, clear differences were observed already on the phylogenetic level of families. An overlap of only 8 OTUs, from more than 100 detected by 16S rRNA gene library analysis, was found between the two vadose pools, and only one OTU was shared between all 3 ponds (together comprising almost 150 OTUs), i.e., including the epiphreatic pool NGIII (Manuscript I, Fig. 3 A). Some deeply branching phylogenetic groups were present only in a single system, e.g., members of the Candidate Phylum OP3 was found exclusively in pool JI. The presence of this group was accompanied by a relatively high number of sequences affiliated with the *Nitrospirae* phylum, and by enhanced nitrate concentration in comparison to the other pools. Both these lineages (OP3 and *Nitrospirae*) were simultaneously observed in several environments including karst water (Alfreider et al 2002, Hugenholtz et al 1998, Pronk et al 2009, Zhou et al 2007); in some cases this co-appearance was directly associated with ammonium oxidation (Egli et al 2003). Thus it might be speculated that there is a connection between water chemistry and microbial community composition in vadose pools. However, since it is difficult to estimate the level of metabolic activity of bacteria in such habitats due to their extremely low cell abundances, it is not clear whether they represent systems with very slowly growing bacteria or if they are only collectors of a set of dormant bacteria originating from diverse niches across the pathway of fed water. In any case, the pools allowed access to microbial community diversity from exotic and poorly studied habitats.

By contrast, the epiphreatic pools, periodically flooded by water originating from conduit aquifers, had similar water chemistry and higher bacterial abundances than in the vadose ponds. A high proportion of the microbiological community in these ponds observed at different time points was composed of the same set of phylogenetic lineages, which were recurrently found during two winter seasons (Manuscripts I, III & IV). Although the number of OTUs shared between 8 samples from epiphreatic pools, as analyzed by Next Generation

Sequencing (NGS), amounted to only 74 of a total of 4804, these shared OTUs comprised more than half of all obtained sequences (Manuscript IV). Different relatedness indicators calculated for the microbial communities of the pools moreover reflected high resilience of some lineages and indicated a considerable stability of the studied assemblages.

4.2 Reproducible trends observed during water residence in epiphreatic pools

Reproducible patterns of the behavior of chemical and microbiological parameters in ponds after flooding events were observed (Manuscripts III, IV). Microbial abundances determined by flow cytometry (Del Giorgio et al 1996) showed initial growth of the bacterial community after flood events followed by a subsequent decline. This was accompanied by decreasing concentrations of organic carbon, and a concomitant gradual enrichment of other chemical parameters. Phosphorus and nitrogen species displayed an increase in concentration only in samples with long residence time (Manuscript III). Our observations suggest that the likely low concentration of assimilable organic carbon (Amon and Benner 1996, Culver 1985) resulted in increasing bottom-up control of the bacterial community and was responsible for the decrease of bacterial numbers. This hypothesis is supported by changes found in DOC composition over time, in particular by the increase of the DOC aromaticity index (AI) (Koch and Dittmar 2006) and of the relative concentration of recalcitrant DOC classes such as combustion-derived polycyclic aromatics (PCAs), soil-derived polyphenols and PCAs with aliphatic chains (Manuscript IV). Some successional patterns were also observed in OTU abundances of different bacterial groups. Although one could distinguish between early, intermediate and late samples based on the composition of β-*Proteobacteria*, analysis of γ-*Proteobacteria* and some *Bacteroidetes* OTUs only allowed a separation of an early and a late group of samples (Manuscript IV, Fig. 4). It should be noted that these community level separation patterns were generated by the different (and sometimes contrasting) temporal development of individual bacterial OTUs, e.g. by subsets of OTUs that increased or decreased over time. At the same time the various responses of OTUs are indicative of reproducible processes in the epiphreatic pools which simultaneously acted on the community during residence.

A decrease in richness and diversity estimators and of mean community distance (MCD) related to the time period that had passed after floods pointed at the selective character of such processes, i.e., the extinction of some lineages from the microbial inoculum and the preferred growth of others.

During the floods or shortly after flooding events, there was a general trend towards low bacterial numbers without connection to the water chemistry. This might probably be a result of the acting of shearing forces on the microbial community during the transport (Farnleitner et al 2005, Pronk et al 2009, Wilhartitz et al 2009).

4.3 Community structure

The influence of deterministic processes in epiphreatic water was confirmed by statistical analysis of the community structure. The observed microbial assemblages displayed a phylogenetically strongly clustered structure as reflected by high nearest taxon indices (NTI) (Webb et al 2002) of all analyzed samples including the one obtained during the flood, representing the unchanged inoculum of karst water. This latter finding was indeed surprising, because the bacterioplankton of streams was proposed to be determined mostly by mass effects, in contrast to water bodies with long residence time (Crump et al 2007, Holyoak et al 2005, Logue and Lindström 2008). Our finding suggests that the bacterial inoculum during the flooding events was probably composed from sets of microbial assemblages already preshaped by selection that originated from habitats through which the flood water was passing by. In addition the transport through the karst environment in itself might be a selective component. The beta net relatedness indices (βNRI) (Webb et al 2002) calculated for the flood water sample and all other samples were not significant, i.e., displayed higher differences of the more deeply branching phylogenetic structure between flood and stagnant water samples. This supports the hypothesis that the inoculum formation involved several habitats with very different properties (e.g., with respect to the presence of oxygen), each with a dominance of characteristic phylogenetic lineages.

4.4 Core microbial community in karst water

The subset of the microbial assemblages present in all samples from the studied epiphreatic pools was conservatively designated as the ”core” community. The largest proportion (in OTUs numbers, as well as in abundance of sequences) of this core was constituted by members of the *Comamonadaceae* family. The majority of the most closely related sequences of these OTUs (in the SILVA database) was obtained from freshwater environments. However, the exact phylogenetic position of most OTUs of *Comamonadaceae* could not be related to any defined genera and had approximately the same distance to *Limnohabitans curvus* (Hahn et al 2010) and *Curvibacter delicatus* (Ding and Yokota 2004). This relatively vague positioning and the rather low resolution of 16S rRNA-based phylogeny

for this group hampers their further ecological characterization, e.g., it is an obstacle for the development of suitably specific FISH probes for the discrimination of these bacteria *in situ* (Newton et al 2011). Interestingly, one OTU, which was shared between vadose and epiphreatic pools and was previously believed to be affiliated with *Rhodoferax* sp. (Manuscript I), also belongs to this badly resolvable phylogenetic group. The closest relatives of 25 abundant core OTUs affiliated with *Comamonadaceae* originated from a karst water study conducted in Swiss Alps (Pronk et al 2009). These facts, together with the high diversification within this lineage hint at the possible ecological importance of this clade in the aquatic karst environment at least within the same geographic region. At the same time, the origin of surface water-borne bacteria in aquatic karst systems remains unclear. A possible differentiation of karst water associated *Comamonadaceae* with respect to habitat should be addressed, e.g. by using genetic markers with higher resolution potential such as 23S rRNA or common functional genes (Ludwig and Schleifer 1999).

A high relatedness with freshwater bacteria was also observed for other core OTUs affiliated with the families *Methylophilaceae*, *Flavobacteriaceae*, *Cytophagaceae* and *Sphingobacteriaceae*, as well as the AcI clade of *Actinobacteria* (Newton et al 2011, Salcher et al 2011, Warnecke et al 2005). This indicates that microbial community metabolism in both environments probably involves the degradation of terrestrial DOC, the turnover of C1 components, and the consumption of amino sugars related to cell wall disintegration (Chen et al 2009, Eckert et al 2013, Nercessian et al 2005).

In addition, microorganisms typically detected in soil were also represented in the core community. Although some of them, like members of *Rhizobiales* (2 OTUs) and *Acidimicrobium* (1 OTU) were less often observed in water samples with longer residence time, OTUs affiliated with the genus *Pseudospirillum* showed increasing abundances with time. This trend was also detected by a prior analysis of two clone libraries (Manuscript III). Unfortunately, since the physiology of this lineage is very poorly described, it is difficult to discuss its possible role in aquatic karst environment.

Although a direct comparison between the data obtained from clone libraries and NGS analysis was not possible because of non-overlapping fragments of the analyzed 16S rRNA genes, the same phylogenetic groups were identified as key players in epiphreatic systems by both methods during two different winter seasons with the exception of members of α-*Proteobacteria* and *Actinobacteria*, which were detected in the second winter only by NGS.

4.5 Abundant but not persistent bacterial groups in epiphreatic pools

Some phylogenetic lineages were often detected in epiphreatic water but did not persist as assessed by our sequencing approach. These groups were frequently related to habitats interconnected with karst. Although it was almost impossible to reconstruct the original communities from the pattern detected in flood water, it is possible to speculate about the origin of a number of bacterial taxa. As an example, 401 very small OTUs (i.e., represented by only 1 or a few sequences) were affiliated with *Myxococcales* of δ-*Proteobacteria*. This lineage is abundant in soils and is characterized by its ability of cyst formation, which explains the continuous presence of these bacteria (albeit not of individual OTUs) during the observed period in pools. Interestingly, high distances were observed between the *Myxococcales* in our samples and their closest neighbors in the SILVA database. This confirms the high potential for novel diversity within this group already detected in some soil habitats (Spain et al 2009). Comparable numbers and sizes of OTUs were obtained for other typical soil-borne bacteria, *Rhizobiales* of α-*Proteobacteria*, whose relative abundances were also decreasing in samples with prolonged residence time. The same trend was also found for some bacterial groups originating from more extreme freshwater habitats, e.g., by OTUs affiliated with the genus *Crenothrix* of γ-*Proteobacteria*. The closest relatives of these OTUs were found in anoxic waters that were rich in methane (Gracas et al 2011). Thus, it is likely that these bacteria were probably inoculated into the epiphreatic pools from an anoxic habitat and that they were not able to adjust their metabolism to the new environment.

4.6 Outlook

In my first investigation (Manuscript I) I noticed a possible connection between water chemistry and microbial community composition in the vadose pools of the Bärenschacht cave system. This preliminary finding would require further investigation. Thus, I am currently involved in a follow-up study to comparatively analyze the microbial diversity in the context of water chemistry in 17 vadose pools from 7 different cave systems. This project is currently in the stage of data analysis.

Another interesting outcome of my investigations in the epiphreatic pools (Manuscript IV) was the tight connection of the detected core OTUs of the endokarst microbial community with phylotypes from freshwater habitats. This relationship might be addressed by future research with additional focus on the origin of these microbes in karst water and their reappearance in interconnected surface water habitats (e.g., Lake Thun in the case of the

Bärenschacht system). In this context, a possible environment-specific micro-diversification could also be analyzed.

Finally, there were specific hints that transport by flood streams itself might have a differential influence on the survival of microorganisms (Manuscripts I and III), i.e. flood water samples displayed relatively low bacterial abundances and levels of bacterial activity. Although other researchers have report similar findings (Farnleitner et al 2005, Pronk et al 2009, Wilhartitz et al 2009), there are currently no studies that asses the direct impact of, e.g., shearing forces on endokarst bacteria. Caves streams would provide a suitable environments to address this question in future studies.

5. References

Alfreider A, Vogt C, Babel W (2002). Microbial diversity in an in situ reactor system treating monochlorobenzene contaminated groundwater as revealed by 16S ribosomal DNA analysis. *Systematic and Applied Microbiology* **25:** 232–240.

Amon R, Benner R (1996). Bacterial utilization of different size classes of dissolved organic matter. *Limnology and Oceanography* **41:** 41–51.

Bakalowicz M, Mangin A (1980). L'aquifere karstque. Sa definition, ses characteristiques et son identificatin. *Mémoires de la Société de France* **11:** 71-79.

Bakalowicz M (2003). Natural organic carbon in groundwater. *Comptes Rendus Geosciences* **335:** 423–424.

Baker A, Barnes WL, Smart PL (1997). Variations in the discharge and organic matter content of stalagmite drip waters in Lower Cave, Bristol. *Hydrol Process* **11:** 1541-1555.

Banks ED, Taylor NM, Gulley J, Lubbers BR, Giarrizo JG, Bullen HA *et al* (2010). Bacterial Calcium Carbonate Precipitation in Cave Environments: A Function of Calcium Homeostasis. *Geomicrobiology Journal* **27:** 444-454.

Barton HA, Spear JR, Pace NR (2001). Microbial life in the underworld: Biogenicity in secondary mineral formations. *Geomicrobiology Journal* **18:** 359-368.

Barton HA, Luiszer F (2005). Microbial metabolic structure in a sulfidic cave hot spring: Potential mechanisms of biospeleogenesis. *J Cave Karst Stud* **67:** 28–38.

Barton HA, Northup DE (2007). Geomicrobiology in cave environments: Past, current and future perspectives. *J Cave Karst Stud* **69:** 163-178.

Baskar S, Baskar R, Thorseth IH, Ovreas L, Pedersen RB (2012). Microbially Induced Iron Precipitation Associated with a Neutrophilic Spring at Borra Caves, Vishakhapatnam, India. *Astrobiology* **12:** 327-346.

Batiot C (2002). Étude expérimentale du cycle du carbone en régions karstiques. Apport du carbone organique et du carbone minéral à la connaissance hydrogéologique des systèmes. Site expérimental de Vaucluse, Larzac, Jura, Région Nord-Montpelliéraine, Nerja (Espagne). Thèse de Doctorat thesis, Université d'Avignon.

Batiot C, Emblanch C, Blavoux B (2003). Total Organic Carbon (TOC) and magnesium (Mg^{2+}): two complementary tracers of residence time in karstic systems. *C R Geosci* **335:** 205-214.

Bekele A, Kellman L, Beltrami H (2007). Soil Profile CO_2 concentrations in forested and clear cut sites in Nova Scotia, Canada. *Forest Ecology and Management* **242:** 587-597.

Boquet E, Boronat A, Ramos Cormenzana A (1973). Production of calcite (calcium carbonate) crystals by soil bacteria is a general phenomenon. *Nature* **246:** 527-529.

Bosak T, Newman DK (2005). Microbial kinetic controls on calcite morphology in supersaturated solutions. *Journal of Sedimentary Research* **75:** 190-199.

Boston PJ, Hose LD, Northup DE, Spilde MN (2006). The microbial communities of sulfur caves: A newly appreciated geologically driven system on Earth and potential model for Mars. In: Harmon RS, Wicks CM (eds). *Perspectives on Karst Geomorphology, Hydrology, and Geochemistry*. Geological Soc Amer Inc: Boulder. pp 331-344.

Boyce JS, Muir J, Edwards AP, Seim EC, Olson RA (1976). Geologic nitrogen in Pleistocene loess of Nebraska. *J Environ Qual* **5:** 93-96.

Braissant O, Cailleau G, Dupraz C, Verrecchia AP (2003). Bacterially induced mineralization of calcium carbonate in terrestrial environments: The role of exopolysaccharides and amino acids. *Journal of Sedimentary Research* **73:** 485-490.

Brooks PD, Mcknight DM, Bencala KE (1999). The relationship between soil heterotrophic activity, soil dissolved organic carbon (DOC) leachate, and catchment-scale DOC export in headwater catchments. *Water Resources Research* **35:** 1895–1902.

Butscher C, Auckenthaler A, Scheidler S, Huggenberger P (2011). Validation of a Numerical Indicator of Microbial Contamination for Karst Springs. *Ground Water* **49:** 66-76.

Cacchio P, Contento R, Ercole C, Cappuccio G, Martinez MP, Lepidi A (2004). Involvement of microorganisms in the formation of carbonate speleothems in the Cervo Cave (L'Aquila-Italy). *Geomicrobiology Journal* **21:** 497-509.

Cai XJ, Lytton J (2004). The cation/Ca^{2+} exchanger superfamily: Phylogenetic analysis and structural implications. *Molecular Biology and Evolution* **21:** 1692-1703.

Cañaveras JC, Sanchez-Moral S, Soler V, Saiz-Jimenez C (2001). Microorganisms and microbially induced fabrics in cave walls. *Geomicrobiology Journal* **18:** 223-240.

Castanier S, Le Métayer-Levrel G, Perthuisot J-P (1999). Ca-carbonates precipitation and limestone genesis — the microbiogeologist point of view. *Sedimentary Geology* **126:** 9-23.

Chen Y, Wu L, Boden R, Hillebrand A, Kumaresan D, Moussard H *et al* (2009). Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME Journal* **3:** 1093-1104.

Clark ID, Fritz P (1997). Environmental isotopes in hydrogeology. Lewis Publishers: New York. p 311.

COST-65 (1995). Hydrological aspects of groundwater protection in karstic area, final report EUR 16547. Directorate-General Science, Research and Development, Europian Comission, Ofice for Official Publications of European Communities: Luxemburg. p 446p.

Crump B, Peranteau C, beckingham B, Cornwell J (2007). Respiratory succession and community succession of bacterioplankton in seasonally anoxic estuarine waters. *Applied and Environmental Microbiology* **73:** 6802–6810.

Cuezva S, Sanchez-Moral S, Saiz-Jimenez C, Carlos Canaveras J (2009). Microbial Communities and Associated Mineral Fabrics in Altamira Cave, Spain. *International Journal of Speleology* **38:** 83-92.

Culver DC (1985). Trophic relationships in aquatic cave environments. *Stygologia* **1:** 43-53.

Davidson EA, Savage KE, Trumbore SE, Borken W (2007). Vertical partitioning of CO2 production within a temperate forest soil (vol 12, pg 944, 2007). *Glob Change Biol* **13:** 922-922.

Del Giorgio PA, Bird DF, Prairie YT, Planas D (1996). Flow cytometric determination of bacterial abundance in lake plankton with the green nucleic acid stain SYTO 13. *Limnology and Oceanography* **41:** 783–789.

Derbyshire E (ed) (1976) *Geomorphology and Climate*. John Wiley & Sons Ltd: London, 524pp.

Ding L, Yokota A (2004). Proposals of *Curvibacter gracilis* gen. nov., sp. nov. and *Herbaspirillum putei* sp. nov. for bacterial strains isolated from well water and reclassification of [*Pseudomonas*] *huttiensis*, [*Pseudomonas*] *lanceolata*, [*Aquaspirillum*] *delicatum* and [*Aquaspirillum*] *autotrophicum* as *Herbaspirillum huttiense* comb. nov., *Curvibacter lanceolatus* comb. nov., *Curvibacter delicatus* comb. nov. and *Herbaspirillum autotrophicum* comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **54:** 2223-2230.

Douglas S, Beveridge TJ (1998). Mineral formation by bacteria in natural microbial communities. *FEMS Microbiology Ecology* **26:** 79-88.

Drake VM, Bauder JW (2005). Ground water nitrate-nitrogen trends in relation to urban development, Helena, Montana, 1971–2003. *Ground Water Monitoring & Remediation* **25:** 118-130.

Drever JI (1997). *The Geochemistry of Natural Waters, Third Edition: Surface and Groundwater Environments*, 3 edn. Prentice Hall: Upper Saddle River, NJ.

Eckert EM, Baumgartner M, Huber IM, Pernthaler J (2013). Grazing resistant freshwater bacteria profit from chitin and cell-wall derived organic carbon. *Environmental Microbiology* **in press**.

Egemeier SJ (1981). Cavern development by thermal waters. *NSS Bulletin* **43:** 31-51.

Egli K, Bosshard F, Werlen C, Lais P, Siegrist H, Zehnder AJ *et al* (2003). Microbial composition and structure of a rotating biological contactor biofilm treating ammonium-rich wastewater without organic carbon. *Microbial Ecology* **45:** 419–432.

Ehrlich HL (1998). Geomicrobiology: its significance for geology. *Earth-Science Reviews* **45:** 45-60.

Engel A (2010). Microbial Diversity of Cave Ecosystems. In: Barton LL, Mandl M, Loy A (eds). *Geomicrobiology: Molecular and Environmental Perspective*. Springer Netherlands. pp 219-238.

Engel AS, Porter ML, Kinkle BK, Kane TC (2001). Ecological assessment and geological significance of microbial communities from Cesspool Cave, Virginia. *Geomicrobiology Journal* **18:** 259-274.

Engel AS, Lee N, Porter ML, Stern LA, Bennett PC, Wagner M (2003). Filamentous "Epsilonproteobacteria" dominate microbial mats from sulfidic cave springs. *Applied and Environmental Microbiology* **69:** 5503-5511.

Engel AS, Stern LA, Bennett PC (2004). Microbial contributions to cave formation: New insights into sulfuric acid speleogenesis. *Geology* **32:** 369-372.

Engel AS (2007). Observations on the biodiversity of sulfidic karst habitats. *J Cave Karst Stud* **69:** 187-206.

Engel AS, Meisinger DB, Porter ML, Payn RA, Schmid M, Stern LA *et al* (2010). Linking phylogenetic and functional diversity to nutrient spiraling in microbial mats from Lower Kane Cave (USA). *ISME Journal* **4:** 98-110.

Farnleitner AH, Wilhartitz I, Ryzinska G, Kirschner AK, Stadler H, Burtscher MM *et al* (2005). Bacterial dynamics in spring water of alpine karst aquifers indicates the presence of stable autochthonous microbial endokarst communities. *Environmental Microbiology* **7:** 1248–1259.

Feth JH (1966). Nitrogen compounds in natural water: a review. *Water Resources Research* **2:** 41-58.

Ford DC, Williams P (2007). *Karst Hydrogeology and Geomorphology*. Wiley.

Forti P (2002). Biogenic speleothems: an overview. *International Journal of Speleology* **30A:** 39-56.

Freeze RA, Cherry JA (1979). *Groundwater*. Prentice-Hall: New Jersey.

Frund C, Cohen Y (1992). Diurnal cycles of sulfate reduction under oxic conditions in cyanobacterial mats. *Applied and Environmental Microbiology* **58:** 70-77.

Fuhrer GJ, Gilliom RJ, Hamilton PA, Morace JL, Nowell LH, Rinella JF *et al* (1999). The quality of our Nation's water: Nutrients and pesticides. US Geological Survey, Circular 1225. p 82.

Gabrovsek F, Menne B, Dreybrodt W (2000). A model of early evolution of karst conduits affected by subterranean CO2 sources. *Environ Geol* **39:** 531–543.

Gounot AM (1994). Microbial ecology of groundwaters. In: Gibert J, Danielopol DL, Stanford JA (eds). *Groundwater Ecology*. Academic Press: San Diego, CA. pp 189-215.

Gracas DA, Miranda PR, Barauna RA, McCulloch JA, Ghilardi R, Jr., Schneider MP *et al* (2011). Microbial diversity of an anoxic zone of a hydroelectric power station reservoir in Brazilian Amazonia. *Microb Ecol* **62:** 853-861.

Griebler C, Lueders T (2009). Microbial biodiversity in groundwater ecosystems. *Freshwater Biology* **54:** 649–677.

Hahn MW, Kasalicky V, Jezbera J, Brandt U, Jezberova J, Simek K (2010). Limnohabitans curvus gen. nov., sp nov., a planktonic bacterium isolated from a freshwater lake. *International Journal of Systematic and Evolutionary Microbiology* **60:** 1358-1365.

Hammes F, Verstraete W (2002). Key roles of pH and calcium metabolism in microbial carbonate precipitation. *Reviews in Environmental Science and Biotechnology* **1:** 3-7.

Hansen B, Thorling L, Dalgaard T, Erlandsen M (2011). Trend reversal of nitrate in Danish groundwater--a reflection of agricultural practices and nitrogen surpluses since 1950. *Environmental Science & Technology* **45:** 228-234.

Hashimoto S, Tanaka N, Suzuki M, Inoue A, Takizawa H, Kosaka I *et al* (2004). Soil respiration and soil CO_2 concentration in a tropical forest, Thailand. *Journal of Forest Research* **9:** 75-79.

Häuselmann P (2002). Cave genesis and its relationship to surface processes: Investigations in the Siebenhengste region (BE, Switzerland), University of Fribourg, Fribourg.

Hill CA (1990). Sulfuric acid speleogenesis of Carlsbad Cavern and its relationship to hydrocarbons, Delaware Basin, New Mexico and Texas. *American Association of Petroleum Geologists Bulletin* **74:** 1685–1694.

Holyoak M, Leibold MA, Mouquet N, Holt RD, Hoopes MF (2005). *Metacommunities - A framework for large-scale community ecology*. Univ Chicago Press, 1427 E 60 St, Chicago, Il 60637 USA.

Hose LD, Palmer AN, Palmer MV, Northup DE, Boston PJ, DuChene HR (2000). Microbiology and geochemistry in a hydrogen-sulphide-rich karst environment. *Chemical Geology* **169:** 399-423.

Huber R, Rossnagel P, Woese CR, Rachel R, Langworthy TA, Stetter KO (1996). Formation of ammonium from nitrate during chemolithoautotrophic growth of the extremely thermophilic bacterium Ammonifex degensii gen nov sp nov. *Systematic and Applied Microbiology* **19:** 40-49.

Hugenholtz P, Goebel BM, Pace NR (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology* **180:** 4765–4774.

ISSKA (2008). Abschätzung des Karstwasservolumens in der Schweiz – Kurzstudie, gestützt auf tektonische Profile, unveröffentlichter Bericht des Schweizerischen Instituts für Speläologie und Karstforschung im Auftrag des Bundesamts für Umwelt BAFU. p 20.

Jassal R, Black A, Novak M, Morgenstern K, Nesic Z, Gaumont-Guay D (2005). Relationship between soil CO_2 concentrations and forest-floor CO_2 effluxes. *Agric For Meteorol* **130:** 176-192.

Kasama T, Murakami T (2001). The effect of microorganisms on Fe precipitation rates at neutral pH. *Chemical Geology* **180:** 117-128.

Kendall C, McDonnell JJ (1998). *Isotope tracers in catchment hydrology*. Elsevier: Amsterdam.

Klimchouk AB, Hydrogeology ISUCoK, Speleogenesis, Society NS (2000). *Speleogenesis: Evolution of karst aquifers*, vol. 527. National Speleological Society.

Koch BP, Dittmar T (2006). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Communications in Mass Spectrometry* **20:** 926-932.

Köhler B (2009). Soil nitrogen oxide and carbon dioxide emissions from a tropical lowland and montane forest exposed to elevated nitrogen input. doctor thesis, Georg-August-Universität, Göttingen.

Kranjc A: About the name kras (karst) in Slovenia. *The 13th international congress of speleology, Brazilia*. 2001.

Kursar T (1989). Evaluation of soil respiration and soil CO_2 concentration in a lowland moist forest in Panama. *Plant and Soil* **113:** 21-29.

Liu F, Liu C-Q, Wang S-L, Zhu Z-J (2010). Correlations among CO_2, CH_4, and N_2O concentrations in soil profiles in central Guizhou Karst area. *Shengtaixue Zazhi* **29:** 717-723.

Lloyd J, Taylor JA (1994). On the temperature dependence of soil respiration. *Funct Ecol* **8:** 315-323.

Logue JB, Lindström ES (2008). Biogeography of bacterioplankton in inland waters. *Freshwater Reviews* **1:** 99-114.

Ludwig W, Schleifer KH (1999). Phylogeny of Bacteria beyond the 16S rRNA Standard. *ASM News* **65:** 752-757.

Macalady JL, Jones DS, Lyon EH (2007). Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. *Environmental Microbiology* **9:** 1402-1414.

Macalady JL, Dattagupta S, Schaperdoth I, Jones DS, Druschel GK, Eastman D (2008). Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. *ISME Journal* **2:** 590–601.

Mattison RG, Abbiati M, Dando PR, Fitzsimons MF, Pratt SM, Southward AJ *et al* (1998). Chemoautotrophic microbial mats in submarine caves with hydrothermal sulphidic springs at Cape Palinuro, Italy. *Microbial Ecology* **35:** 58-71.

Beck BF (ed): Manganese deposition in limestone caves. *8th International Congress of Speleology*; Western Kentucky

Mueller DK, Helsel DR (1996). *Nutrients in the nation's water: Too much of a good thing?* US Geological Survey: Denver, CO, ETATS-UNIS.

Neff JC, Asner GP (2001). Dissolved organic carbon in terrestrial ecosystems: Synthesis and a model. *Ecosystems* **13:** 29–48.

Nercessian O, Noyes E, Kalyuzhnaya MG, Lidstrom ME, Chistoserdova L (2005). Bacterial populations active in metabolism of C-1 compounds in the sediment of Lake Washington, a freshwater lake. *Applied and Environmental Microbiology* **71:** 6885-6899.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol Rev* **75:** 14-49.

Nishio M (2002). Effect of intensive fertilizer use on groundwater quality. Institute of Agricultural and Forest Engineering, University of Tsukuba, Japan: Tsukuba.

Northup DE, Barns SM, Yu LE, Spilde MN, Schelble RT, Dano KE *et al* (2003). Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environmental Microbiology* **5:** 1071-1086.

Palmer AN (1991). Origin and morphology of limestone caves. *Geological Society of America Bulletin* **103:** 1-21.

Peck SB (1986). Bacterial deposition of iron and manganese oxides in North American caves. *NSS Bulletin* **48:** 26-30.

Plummer LN, Wigley TML, Parkhurst DL (1978). Critical review of the kinetics of calcite dissolution and precipitation. *Abstr Pap Am Chem Soc* **176:** 84-84.

Porter ML (1999). Ecosystem energetics of sulfidic karst. Master thesis, University of Cincinnati, Cincinnati, OH.

Priesnitz K (1974). Lösungsraten und ihre geomorphlogische Relevanz. *Abhandlungen der Wissenschaften in Göttingen, Mathematisch-Physicalische Klasse 3* **Folge 29**.

Pronk M, Goldscheider N, Zopfi J (2006). Dynamics and interaction of organic carbon, turbidity and bacteria in a karst aquifer system. *Hydrogeol J* **14:** 473-484.

Pronk M, Goldscheider N, Zopfi J (2007). Particle-size distribution as indicator for fecal bacteria contamination of drinking water from karst springs. *Environmental Science & Technology* **41:** 8400–8405.

Pronk M, Goldscheider N, Zopfi J (2009). Microbial communities in karst groundwater and their potential use for biomonitoring. *Hydrogeol J* **17:** 37–48.

Raich JW, Schlesinger WH (1992). The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus Ser B-Chem Phys Meteorol* **44:** 81-99.

Reischer GH, Haider JM, Sommer R, Stadler H, Keiblinger KM, Hornek R *et al* (2008). Quantitative microbial faecal source tracking with sampling guided by hydrological catchment dynamics. *Environmental Microbiology* **10:** 2598-2608.

Rhue RD, Harris WG, Nair VD (2006). A retardation-based model for phosphorus transport in sandy soil. *Soil Sci* **171:** 293-304.

Ruby EG, Wirsen CO, Jannasch HW (1981). Chemolithotrophic Sulfur-Oxidizing Bacteria from the Galapagos Rift Hydrothermal Vents. *Applied and Environmental Microbiology* **42:** 317-324.

Rustad LE, Fernandez IJ (1998). Experimental soil warming effects on CO2 and CH4 flux from a low elevation spruce-fir forest soil in Maine, USA. *Glob Change Biol* **4:** 597-605.

Rusznyák A, Akob DM, Nietzsche S, Eusterhues K, Totsche KU, Neu TR *et al* (2012). Calcite Biomineralization by Bacterial Isolates from the Recently Discovered Pristine Karstic Herrenberg Cave. *Applied and Environmental Microbiology* **78:** 1157-1167.

Salcher MM, Pernthaler J, Frater N, Posch T (2011). Vertical and longitudinal distribution patterns of different bacterioplankton populations in a canyon-shaped, deep prealpine lake. *Limnology and Oceanography* **56:** 2027-2039.

Sanchez-Moral S, Portillo MC, Janices I, Cuezva S, Fernandez-Cortes A, Canaveras JC *et al* (2012). The role of microorganisms in the formation of calcitic moonmilk deposits and speleothems in Altamira Cave. *Geomorphology* **139:** 285-292.

Sarbu SM (2000). *Movile Cave: a chemoautotrophically based groundwater ecosystem*.

Simmleit N, Herrmann R (1987). The behavior of hydrophobic, organic micropollutants in different Karst water systems. *Water, Air, & Soil Pollution* **34:** 79–95.

Simon KS, Pipan T, Ohno T, Culver DC (2010). Spatial and temporal patterns in abundance and character of dissolved organic matter in two karst aquifers. *Fundam Appl Limnol* **177:** 81-92.

Spain AM, Krumholz LR, Elshahed MS (2009). Abundance, composition, diversity and novelty of soil Proteobacteria. *Isme Journal* **3:** 992-1000.

Spilde MN, Northup DE, Boston PJ, Schelble RT, Dano KE, Crossey LJ *et al* (2005). Geomicrobiology of cave ferromanganese deposits: A field and laboratory investigation. *Geomicrobiology Journal* **22:** 99-116.

Trček B (2008). Flow and solute transport monitoring in the karst aquifer in SW Slovenia. *Environ Geol* **55:** 269-276.

Tucker ME, Wright V. Paul (1990). *Carbonate sedimentology*. Blackwell scientific publications: Oxford London.

van Beynen PE, Schwarcz HP, Ford DC, Timmins GT (2002). Organic substances in cave drip waters: studies from Marengo Cave, Indiana. *Canadian Journal of Earth Sciences* **39:** 279–284.

van Beynen PE (ed) (2011) *Karst management*. Springer Dordrecht Heidelberg London New York, 489pp.

Warnecke F, Sommaruga R, Sekar R, Hofer JS, Pernthaler J (2005). Abundances, identity, and growth state of actinobacteria in mountain lakes of different UV transparency. *Applied and Environmental Microbiology* **71:** 5551–5559.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002). Phylogenies and community ecology. *Annu Rev Ecol Syst* **33:** 475-505.

WHO (2011). *Guidelines for drinking water quality-recommendations, 4th editon*. World Health Organization: Geneva.

Wilhartitz IC, Kirschner AK, Stadler H, Herndl GJ, Dietzel M, Latal C *et al* (2009). Heterotrophic prokaryotic production in ultraoligotrophic alpine karst aquifers and ecological implications. *FEMS Microbiology Ecology* **68:** 287-299.

Wu Z, Yuan L, Jia N, Wang Y, Sun L (2009). Microbial biomineralization of iron seepage water: Implication for the iron ores formation in intertidal zone of Zhoushan Archipelago, East China Sea. *Geochemical Journal* **43:** 167-177.

Xuan L, Zhijun W, Junbing P, Pingheng Y, Qiufang H, Zhiyong H: Nitrate fate and origin in a typical karst subterranean river system, Southwest China. *5th International Conference on Bioinformatics and Biomedical Engineering* 2011.

Zeng J, Yang H-m, Lou K (2010). Prokaryotic diversity of a non-sulfide, low-salt cold spring sediment of Shawan County, China. *Journal of Basic Microbiology* **50:** 484-493.

Zhang WL, Tian ZX, Zhang N, Li XQ (1996). Nitrate pollution of groundwater in northern China. *Agriculture, Ecosystems & Environment* **59:** 223-231.

Zhou J, Gu Y, Zou C, Mo M (2007). Phylogenetic diversity of bacteria in an earth-cave in Guizhou province, southwest of China. *Journal of Microbiology* **45:** 105–112.

List of publications:

Zeder M, Peter S, Shabarova T, and Pernthaler J (2009) Cultivable planktonic Flavobacteria that are highly active during the spring phytoplankton bloom in Lake Zurich. Environ. Microbiol. 10: 2676-2686

*Shabarova T, Pernthaler J (2009) Investigation of bacterioplankton communities in aquatic karst pools in Bärenschacht cave of Bernese Oberland. Proceedings of 15th International Congress of Speleology, Texas, USA. 416-421

*Shabarova T, Pernthaler J (2010) Karst pools in subsurface environments: collectors of microbial diversity or temporary residence between habitat types. Environ. Microbiol. 12:1061-1074

Schmelcher M, Shabarova T, Eugster M R, Eichenseher F, Tchang, V S, Banz, M, and Loessner, M J (2010) Rapid multiplex detection and differentiation of Listeria cells with fluorescent phage endolysin cell wall binding domains. Appl. Environ. Microbiol. 76:5745-5756.

*Shabarova T, Widmer F, Pernthaler J (2013) Mass effects meet species sorting: transformations of microbial assemblages in epiphreatic subsurface karst water pools. Environ. Microbiol. (In revision)

*Shabarova T, Villiger J, Morenkov O, Niggemann J, Dittmar T, Pernthaler J Bacterial community structure and dissolved organic matter in subsurface karst water pools: interplay between habitat filtering and flooding events. (submitted to the ISME Journal)

Kohler E, Blom J, Ewert C, Villiger J, Shabarova T, Derlon N, Posch T, Morgenroth E, Pernthaler J Degradation of microcystins in a gravity-driven ultra-low pressure ultrafiltration system (In preparation)

Participation of meetings:

Feb. 2009: Oral presentation at the 3rd Swiss Microbial Ecology Meeting, Einsiedeln, Switzerland

Jul. 2009: Oral presentation at the 15th International Congress of Speleology, Kerrville, Texas, USA

Jun. 2010: Poster-presentation at the 69th Annual Assembly of the Swiss Society of Microbiology, ETH Zürich, Switzerland

Aug. 2010: Poster-presentation at the 13th International Symposium on Microbial Ecology, Seattle, USA

Feb. 2011: Oral presentation at the 4th Swiss Microbial Ecology (SME) Meeting, Engelberg, Switzerland

Oct. 2011: Oral presentation at the 8th International Symposium of Subsurface Microbiology, Garmisch-Partenkirchen, Germany

Sept. 2012: Oral presentation at the 13th National Congress of Speleology, Muotathal, Switzerland

Acknowledgments

I would like to thank:

My supervisor, Jakob Pernthaler, for giving me the opportunity to work on this very exotic and interesting topic, for the interesting and stimulating discussions on various scientific and ... non-scientific subjects, for infinite patience and especially for his confidence in me

Franco Widmer for agreeing to be a member of my PhD committee and for his help in T-RFLP setup and data evaluation

Leo Eberl for investing considerable time in the committee meetings and useful remarks

Roland Psenner for providing insight into limnochemistry and Austrian alpine limnology and allowing me to use infrastructure of the Institute of Ecology for water sample analysis

Thorsten Dittmar and Jutta Niggemann for the opportunity to conduct FT-IRC-MS analysis and their hospitality in Oldenburg

Hans-Peter Kohler and members of AuA lab of EAWAG for chemical analysis of water samples

Friedrich Jüttner for the introduction into gas chromatography, sharing with me his broad knowledge, undying optimism and creative ideas

Jörg Villiger for his professional engagement for 454 analysis (sometimes also on weekends), useful discussions and help in mobbing of Stefan

Eugen Loher for taming of laboratory gadgets, my exclusive “microbiological devices” and especially for flow cytometry and gas chromatography assistance

Josef Franzoi for very competent chemical analysis of “unique cave water samples”, and for this incredible understanding of contradictory problems of the Russian mind

Salome Schneider and Roy Kissling for helping me a lot in Franco Widmer’s is lab

Oleg Morenkov for mathematical and statistical support, discussions about nasty analysis, proofreading and his help during difficult moments

Diego Sanz for caving assistance during long sampling campaigns, photos, support and patience

Rolf Siegenthaler and Philipp Häuselmann for their help in the finding of objects of interest and providing important information for sampling organization in Switzerland and Romania

Martin Achtman for his help during sampling campaigns and being my best caving partner ever

Lukas Plan and Austrian cavers for their help in organization and realization of sampling campaigns in Austria, photos and nice time on several expeditions

Ulrich Meyer and all Untersberg caving team for inviting me to join the expeditions in Germany and their help during sampling

Ester Eckert for the best work atmosphere in our office, discussions, proofreading, help in investigation of etymology of German proverbs, friendship and support

Michaela Salcher for interesting discussions, proofreading, sharing not just her scientific library, funny diving experience in Lake Zurich and conversations about Russian politics

Esther Kohler for a very friendly atmosphere and methanol zero-tolerance agreements

Stefan Neuenschwander for help with RAXML algorithm tuning, interesting scientific discussions, proofreading and endless conversations about our next climbing date

Bettina Eugster for sharing with me the early morning hours at the station

Thomas Posch for his care, nicknames, interesting discussions and chocolate

Judith Blom for discussions, exchange of lab and life experience, and proofreading

Marie-Ève Garneau for discussions, especially about Canadian cinema

Markus Steinkellner for his help in the lab and ordering

Astrid Kunz for her administrative support and delicious cakes

All members of Limnological Station Zurich and SGH-Bern